

Figure 1

SEQ ID No: 1

gtctataatggcagtcacacagggctctaaaactttgcagttttatcattaactcaaagtgaatgtatacatg
ccgctgactcaacattttgagagacaacaaatacaatgaatatcaagatacatatataatataatgtat
ctttttgagatggagtttactgtgtgtgtccaggctggagtacaatagcacgatcttggctcactgcaacc
tctgcctcccaggttcaagcaattgtcctgcctcagcctcccaagtagctaggattacaggcatgtgccacc
acacctggctaattttgtattttttaagtagagatgggggtttcacctatgttggtcaggctgggtctcgaactc
ctgacctcaggtgatccacctgcctcagcatctcaaagtgtgggattacagggtgtgagccacccacccgg
ccatataatataatttttgagatggagtttactctgtcaccacaggctggagtgaatggcttgatctcggc
tcattgcaacctctgcctcccagttcagatgattctcctgcctcagcctctcaagaagctgggattacagg
tgcagtcaccacatgcccactaatttttatatttcatcatgggggtttcacctatgttggccagggtgggtgtcg
aactcctgacctcaagtgtctgcctgccttcggcctcccaaagtgtgggattaccggcatgaaccaatac
gcttggcaataattttttaagaaaaaaaatttcaggttgcaacagcatccaaaaagtaaccaatgatttta
gggtgaagggtgaagacaaatgtaaactcttttttttttttgaatggcgtcttgcctgtcgcctaggct
ggagtgcatttgggtgcaattgcgactcactgcaacctccacctcctggactgaaacgattctcctgcctcagc
ctcccagtagtgggattataggctcgcgcgcgcacgccccgctaatttttgtgttttttagtagagacggg
gtttcacctatgttggccaggctgggtctcaaactcttgacttcaagtgtatccgctgccttggcctcccaaag
tgctgggattacaggcttgagccacccgggtgaatgtaaattgttaaactgtgttttgaatgcataag
tataggataagggaattgactttctgaagaccagaacatttttagtcaatttcaaacacaatgtgagtcaa
ttgtataaaacaggttcttctcctgatgaggataagaatagtagtcttctgcagatggaaatgccattcag
ctgtactttctagtgggttacgcccatagtagcactgttgatggaaccaggtatctgactttaggaaagatgt
tccccactggagctgacctcagaggagcctgaccaacttggggaaagttaaagatctcatcacgtggagaa
taggggaaggcaccaacacgtattgagtgtctactttgagcttaagggaaggagaaaaggcagggaataa
acggaggatggaataagaataggtaattctccttaggtttaataataagtgcttgccataggaaggagccc
cagaacacagttatcaataatagagactcacacagagcattctacactagagctgctgtcctcttgaccaga
ataagggttaagggtgtgtgtgcgtgtccaggaaagttaggcagctaggaggtgatcagagcatacactactgcc
gccacaattctaaagtgtcttcccctagggggaatcctatttcttctcaggcacatttgtttattcattccatg
ttcactcttgttatttacttcttggcaggctttgtgttaagaattggggaaacaagggtgaataaaccaggt
ctgtaaagaaaaggagctcacagctcggaggggcaaatgggcattgtgcctgcaagttggccactgagagc
ctaagaagtgaagttatgaatccaggattactcagttatcaatgaagtgattaaacatcatccatacagacc
ttcagagctggagggaattttggatacctactcagcacatagttttcaaacagtgcttgtggaaccctagg
gcatttcttagggattgtgtgtgtgtgagagaggagattgaatcagaagggtgtctgggaccattctctactca
cacttcaagcagagcagctccacttctatctgtattattttttattttttatttttztattttttttga
gacggagtctcgccctgtcgcccaggctggagtgcagtgccacgatctcagctcactgcaacctctgcctcc
cgggttcaagagattctcctgcctcagcctcctgagtgtggaattataggcctatgccaccacgcccagc
taatttttgtattttttagtagagacgggggtttcacctatgttggtcaggctgggtctcgaactcctgacctgt
gatccaccggccttgcctcccaaagtgtgggattacaggcattagccaccatgccccgcctatctgtatt
atttattcattattgtatgtgaatgaacctgaagaatgcttactgttactgctaagattttaaccacaccc
catgcccattgcaggatgatagtgaatagtgggccaaaagatactataattagactcatgtaattaaggaatat
ttttgtcttgtacctattatgtgcctataaagactatgaaatctattttattcagtgattttattggaatacca
aataagcaaagatcctatgtgctaaagattctaatttgtgctaagattttccttcagatgtttggctttct
caaattccctgagggctagaactttgcctactcatttgtgtttcccaagtgtctaacgcagtgctgacac
atacaggatctccaaacgcttgcctgaatgtgtgaggaaggaattaaaataatgtaccgcgggcaagtgcc
tcatgctataatcccagcactttgagagaccgaggtgggcagatcacttgaggtcaggagctcgagcccag
cctggccaacatgggtgaaaccccgctctactaagaatacaaaaattagccgcgagttgtggcaggcgctg
taatcccagctactcggaaggctgaagtgggagaatcacttgaacctgggaagcagaggtcgtagtgagccg
agatcagaccgctgagctccagcctgggtgacagagcaagactccatctcgaaaaataaaagtaaaataaaa
taatgtactaactggaccacagaacagattttccaattgattattgacaacaaaggaatctgaattatttaaat
aagggtgaataagtagcatattcatatataatgtatattgtgtgtgtattttacattttttataaaagtgtaaaag
tatatatacttttttttctttcttcaggtagaacacctcctagattgtcactgaataaacattagcactaa
ctatggcaatcaaatacatattgattgggtgtcagagaagaattgaacattcaactctgaagcagtggtatt
tcttccatctgcaaacactctgtccccatcchtctttgtgtatcctggaatccaagtcataaataatgat
aggatttctgtcaaaggattttctcagaggagtgatattgtaactcccttctcctctgatacactgactcact
aagcagctacccttgtgaattccaattagcaatacttcttgcctatgtctgggtccaactttcagacaaacct
gtgttcaggattcctatagccatttatagggtgtagacaggaagcattcagatatccccagagggtacctgata
accagctgatccatgactgtctgtcttgggcttggccagcttgaaatcttgacattgtgggttctczzccaga
gaagggtgccttttggatgtgagataaagacattatgactagatagtgcatgggtgaggggtgtttttctagttt

[illegible]

gtcccagctactcgggaggctgaggcaggagaatggcatgaacccaggaggcgagcttgacagtgagctga
gatcacaccactcgcgtccagcctgggccacagagcaagactccatctcaaaaaaaaaaaaaatcttttata
ctatattcttaatgtaccttttctatgttttagatacacaaataccactgtattacaattgcctactgtattc
agtacagtaacatgctatatggttacgtagcctaggaacaataggctatatgttccctagggtatagggatgt
ggatccctataccatctagggtttgtgtaagtttattttatgatgtttgcatgacgacagagtcacctaagga
ctcattttcttagaataatattttagttagtaagcaatgcatgactctattgactcatgaattcttaccacagacc
tatggggcagtactattgttaccctcattttataaatgataaaactgaggtacagagacagtaaaataacttg
accacggtcattcagctactcaagagtcaaggctgggattttaaaaccagatcacatgggttcagagtggtca
cacttacctactatactgtctcaagagcaaggatgttttggttcacttgacaaatgaagatagggaacctctt
tcattataagcctatttttaggctaaaaatagaagggaaggggacacagtgaaatccaggcctctggcagtggt
cctcagccctttctgagctggcctgggacagccttccactcactgatgccacttccactgagcgacttt
cctgcctacctcactgatgccacttccactgagcgactttcctgggctccagacccagtaagcgactttgc
ctgcaaccaccttattttgcctactccctgtgcttttatgcctttaccatctgccctggaaagctcttct
aacctttgaatgggtaaaggcataaatgtatgctggagaaatcctcagctcagggccaggcacgctgggtca
cgctgtaatcccagcactttgggaggccgaggtgggtagatcacctgaggtcgggagtgaggagaccagcct
gaccaacatggagaaacccctgtctctactaaaaatacaaaattagctgggcgtgatggcacatgctgttaa
tcccacagctactcgggaggctgaggcaggagaattgcttgaacctgggaggcgagggtgagctgagccaa
gattgcgctctgcactccagcctgggcgacagagctagactctgcctcaaaaaaaaaaaaaaagaa
aaaaaaaaaagaaaaaagaaatcctcagctgagttgtcaactcctctttgaaactttctcagaccttccaggc
tgagtcgctcgtcattttgtgcttccctcagttccctggcttctaccttcttcatagcttgtttcatgtaattgt
aattcttacttgcctttctccctcttctaagctgagagctacttcaaagcatgggtaggacctagcacggtgt
atgggacatgggtggtaccccgtaaatgtttactgaaaaaaaaatgcctaaagcaattgttaacatcatcag
atagataattatgggcattcagagattctgtcttcaagcttatataaagaacttatttttggctctaattat
cctgataattttctcattactttcacttattgtggcttgtggatcaattgttgacattttataaacatttca
ctatttgacaatgatgatactaaaaatacgaattaaagcaaccttctaaagatagtgtatgataaacatata
cgctggtaaacatctttattttcagccgtatcatggaatcctctgtttccattctgctaggtaggcaggtatg
caggtagaacttgtgagaggatattgatttttgtttccatcttagatatgacaggaacttggaaattttgaca
taaatgacgaacatccgggattcttaacaatctttaaaaatggaatgccttaaaagctgggcgcagtggt
cacgcttataatcccagcactttgggaggctgaggcaaatggatcacttgagttcaggagttcaagaccagc
ctggccaacatggtgaaaccccatctctactacaaatacaaatattagccgggcgtagtggcagggcgctgt
aatcccagctacttgggagcctgaggcaggggaattgcttgaaccaggaggccttggagattgagctgagc
tgagactgcgccattgcactccagcctgggcaacaagagtgaactccatctccggaaaaaaaaaaaaaaaa
aaaaggaatgcctttgggaataatattttataatttatgtataacatatagacaaaccattagtttgtctt
atattttactaaatataaatatttagtaaatataaatatttactaaatataaaaaactcttagattttactaaag
agttacaactaattggcctggcgtgggtggctcacacctgtaatcccagcactttaggaggcagaggtgggccc
gatcacgaggtcaggagatcaagaccatcctggctaacacggtgaaactctgtctctactaaaaaaaaaaaa
atacaaaaaattagccgggcgtgggtggcagggccctgtagtccagctactcaggaggctgaggcaggagaa
tggcgtgaactaagcagaggttccctaaaagtgtcttcaggataaaggcagaggaagaggctccatgact
gggattgggtgtgaggagagccagagaagcaagctacagaaaaagagaaaaaattaatatgcaagagagtaaa
caacacgaaggtaaaagaaacccaggtgtggaacactacagctgagaaagggtgtctgaaggatgtctacaaa
gcaaatgcttggatattaattcattgacagcaggagatggtaagcctcatgataaagaaggagaaaaaatcaa
gtcaagggtctgaggtactgacccagggtatacttgactatgccagcaactgtttagggggagatttgagct
acactttagcaaaaggcaaaatctgtaattagttgtaactcttttttttttttgagatgggtgtctcgctctgt
ccccaggctggagtgagtggtgtgatcttggctcactgcaagctccgcctcctgggttcaagtgtattctc
cagcctcagcctcccaagtaattgggactacaggcatgcaccaccatgccagctaattttttgtacttttat
tacagaccatgttttgccatgttcaccaagctgggtctcaagctcctgacctcaagtgtccgtccgcctcgg
ccztcccaagcgctgagatttataggcctgaaccaccgcgcctggcctaaagagatctaattcttagcaaag
tttcaccagggtgtctctcctcaccocccaccccatccttcccacaaagaattagaacaatgtccctactacc
cctgctgtatctctgactttttactttaaatctcagcagaatattttactaaatgttttgatgtgggtatat
aaaatcatccctgctgacaaggaaacactttttgaaaaaagttttcattatcaaacagtaagtacagctgac
tgccgtgacctttaaccattttctgagctctccctcattggacttgggtggaggggactgggtaccaataaag
tcaaatgcttaataatttatgcaagtgttgaagaaatttgaagttgaatatttctatcatcttgaatgga
gaaagaatctgtaaacagcaaaagccagacgcctaaaggaaaagatttacagattaaaataagattgcaatc
tggttaaaaaattttgcaacacatgtaacagtgagaaagttgaaacacttggtttaacawgagcttttawcag
ataaataaggaaagaataaacatttggattttaaacaccgataaacatgaaaagatgttttaactctttattttt
atttaatccatattatttttcagtttaatacaagaaaataataatcaaaacaataataacattttatatatat
atatatatatatatatatatatatatatatatagtaggcataaggttaagactgataagactgttgaaaagg

gatgaaaaaactaggcttactcatcaccaatatatatctatttaggttggtctaagtaaaaaaactgaaaaatat
caagtgtctaaaaggatattggagcaattggaaccctcagacatcgctgtrtgagaaaaacaaaatggtacagc
caccctggagaacagtttagctgtttcttgtaaagttaaacatgcgcttaccatatgactcagcaatctcac
tcttgggtattttatgctaggaaaaggaaaaatttatacttgcacacaaaaaacttgtaaagtgaatctttatag
cagctctattcataactgccaaaaactgagagaaaaatgtcctttaatgtgtgaatggataaaccaactgtgc
aacatccatgtaatgaaatactacttagcaataataataataattataaaaaccccagaaccattgatgcat
gcaacaaatatggataaatctcaaaagcattatgctgagtaaaagaagtgcgtctgaaggatttcatactct
aggattccatttatataacattattgaaacgacaaaattatggggacagagaatagatcagcgggtgccagg
ggtttagggtgtgtggagagggtgtggctataaagaacatgcaaggggaattttttggggagatgaaatggatc
tgtatcctaattatgggtcatggtaacacaaaatctatacatgtgttttagattcatagaactgtataccaaaa
aaaaaaagtcatttttactctcttaaaatgaaaaaagaaaaagcctgggcattctaacaccttgtttgtgag
agtacacattgataccaagttttatgggtgggcaattttgtctataaatacggaaagtttgtctgttctattat
tcagcaatcccagttttgcaaaaactgtcataagaattcttggggcgccgcagcgggtgttcacgcctgtaa
tctcagcactttgggaggccgaggtgacgatcactgaggtcaggagtttgagaccagcctggccaacatt
gggcaaccctgtctctactgaaaaatacaaaaatttagccgggcatgggtggcgcatgcctgtagtccgagctac
tcgggaggctgaggcaggagaatcacttgaacctgggaggcagaggttgtagtgaactgagatcgtgccacc
gcactccagcctgggcaacagagtgagaatccgtctcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacaaaaa
aaaaactttgtgtacgtgtgcaaagagaatacaaatgatcatggctgcattttttaaatgactataaaaa
agaggtacaaccagccaggtaaagtgggtgtgcacctgtagtccagctactcgggagggtgaggtgagagga
acacttgagtccaggagtttcaggccagcctgggcaacatagtgcagccctgtcccaaaaacaaaaa
caccaaatgtctatctgttaggaatttggtttcaagttgtgatcataggtacagtgaatatattacattc
atttaaaatgatgataaaatctgtattttgtttacatgaaaaactgtccactataatgtagtgaataatag
attacaacaatatatatggaataaacttgttttagaacaattttctagaagaaggtaaatggacagaatta
tctctgggaagtgggtttataatgattctcattttcttctttgtatctttttcatagtctttctacttttgt
tatgtctggacatttgattatgagcatgtattactgatctattttaaaaaattgattttaatttttacaaaa
ctcatgtaaaagcttgaaggttcgcatttttagaccatgttaaaattttctctggatcaaaacagacttattcaa
atatcttgtaacctgtctctcaaaaattgcctgccaaaatacactataaaagagagcatttagctgcatttt
ttggactgtctgatgatacaaatattatttaccatggcttaattttttaccctccagatgtgtgtgggtaca
aacactctccacattttttaggcatttgcttttgatatttttaaatgtaaatcagctgtgcgctgggtgta
cgctataatcccagcactttgggacgctgaggaaaggatcacttgaggtcaggagtttgagaccagtttag
ctaactgggtgaaaccccgctctcttttaaaactacaaaaatttaaccgggcatgggtggcaggcacctgtaac
ccagctactcaggaggctgaggcaggagaatcacttgaacttgggagacagaggttgcagtgcgagatc
atgccactgcctccagcctggccacagagcgacactccatctcaaaaaaaaaaaaaaaaaaaaaaaaaaag
gccaggcgagtggtcacgcctgtaatcccagcactttkggagcgccaaggtggssggatcaactgaggtt
gggagttcacgaccagcctgaccaacatgcagaaaccygtctctactaaaaatacaaaattagccgggtgt
ggtggtacatgcctgtaatcccagctactcgggaggctgaggcaggagaattgcttgaaccaggaggtgaa
ggttgtgttgagctgagatcccgccattgcactccagcctgggcaacgagcaaaaactctgtctcaaaaaaac
cgaaaaaattcccccaaaaaacaaaaaaaacagcaacaacaaaaaatacaataatgtaccttgttttagc
ataaagcataattatatgcatatgggtgattgggaggatgaaatggaaagggttatttattactgacttcagaa
attatgtcctgatagatttgattgggtgatttaaatataacttcttgtcaagcatctgtttttagaatcaaat
actatgactctgcagtttcttgaactctcatgatatcacatctctgtttgcctttgcattgggttttaagaaa
atgaggagtggtgaaaacttcaacttcgtttcatgtattacatttttgaatgacacactggctatttctcta
gaaatataagggtgacaaatatttcacagaaacataagggtgctattatctcattcaatcttaggtcactcaa
actctttctctccacacattgaagattcatttgggaatgcttttgtcttatttgtgcacccccagtgagggt
gtggtaagtgtttttcattttgtcttctttgttttatctacagggttccattcaataaaacaaagggtggtg
tcaaacttcaggctcttatgggtttggatgtaatctttgggtctcatttttagttaccaacagagaggtgtgc
ttctgacctctttgactcttccctgtgtaatttactatgcctttgatacttgtgaagggtgagattttcgag
gagtactgttgtttttgttagaggttgtaatgtctttcttgcctttgtgattcaagttgtgttcagttacaa
tcataagcatgtgcctaaaaaaatcagatgcaaaactagcaaaagtagaaactcaggggtgacagctcttaaga
aaagatgcaattcttggggctgggtgcggtggcttatgcctataatcccagcatgttgggaggccaaagggtg
gcagatcgcccagggtcaggagttcgagaccagcctggccaatgtggtgaaacctgtctctactaaaaata
caaaaattagctgggcgtgctgggtgggtgctgtaatcccagctactcaggagggtgaggcaggagaattgc
tggaaacctgggagggtggaggttgcagtgcgcttagattgcgccattgcactccagcctgggcaacaagagcg
aaactctgtctcaaaaaaaaaaaaaaaaaaaaaaagatgcaactcttattactgacacagaaatgaaaaattag
ttacatagatttgtaaaaggactatcagtaggttttagccttaccagatttaggtaaatcatttctgtcta
cactcatattctcagccacttccctcatcacattttcagggtgcagtatataatagcgtcaactcgtgtaatt
tccccctactccccatgaacttctagqccaaqqqqccacacqqqqttqqqqcatataqtataaaggagtaaqg

[illegible]

agagtttctagcaaccagtatcatttggtatttttaacaatgtgtacatgtacatttatgcagatgagttaac
atatatcaaagcaacctccaaacaatgccatttaggtaatctccaatttaaagcctcaatagaatgataag
attgagcttttctgtagttccatgacctccagcagagtctgcaaggccacagctgcctgaagggttgattctg
taattagaagatgccaggggtcatctcagaatagaacctcaagccacccaggctacatttacagaatcagcct
ctccagaaaaacagcaacaaaggagggccttccatgtatttggaaggagtcacctagaggagggacttggg
gttttggtgttggtgtggggggcagggatgggatggggagggggaagcttattgaaatatactaaaagacaa
accaacctaaagggtggaggggaagaaaattcacacttgtaagcttcttttttaaggggcatctcttaggctc
tagcttttgagattcagtatatatatttttgagtcttgctctgctggagtgcagtgggtgtgatctgggct
cactgcaagttctgcctctcaggttcacaccattcttgctcagcctcccagcagctgggactacagggc
gcctgccaccatgcccagctaattttttgtatttttagtagagacggggttcaccctgttagccaggatgg
tctcgatctcctgatctcatgatccgcctcgccctcgccctcccaaatgctgggattacagggcgtgagccacc
gcgcccggcccagtatttttgttttatgaagatattacatttgtaagtatgagcttggtgtcagcaaacctt
atatccctgtgtacaaactggccgaatcacttagccactttgggccaatcacttagctcttctaacagtaa
gaaatcaacaagaaaaataaacatttcaaacattacaatgtgttcattgattcactgtgggggatgaccaga
ttctcgaaccacaggttgttcttagtgaaacaagttgggttggggccatagacttgtgtatttagaatca
atggctgttctctctctgggactttgattttttcttgggctcatccctttttgtagtatcttattcttgt
cttatttgtataggacttaactgttccattcccttatttagagcaatctaagtgattaccttcataaccttt
ggaattatatgcttcaaaattccaaaaagaatgattttgggctgggcacagtggtcacacctataattcca
gcactttgagaggctgaggtggatcgctgaggtcaggagtttgasrccawcctggccaacgatagtgaac
ccgctctctactaaaaaatataaaaaattagccgggcatggtggcaggtgcctgtaatctcagctactcggg
aggtggaggttgcaatgagcccagatcgccaccattgcactccagcctgggcaacaaaagggtgaaactccatc
tcaattaaaaaaaaaataatgattttgggtggtgcacttcaaataggtaggagaagaaggagagaggagatgg
agggctcasggagatctaattactctctaaaatcatgctaggaagataaacaccttttaataacactctctgc
ttttataacatcattctgccaaggagctcaaaggtttcaacamagttcactttcagaaaacccctttgagga
agacagaatatacatcttctctccmtttttaagatgaagaaacaggccgggcacaaatggctaagtgcctgtaa
tcccagactttkggaggctgaggccasargatcgcttgagctccaragtttgagaccagcctggataacat
ggcaaaacccctgtctctacaaaaaaaatacgaataatagatgggtgtgggtggcatgcacctgtggtcccagc
tacttgggaggctaaagtgaggatcgcttgagcccaggagtcaggtctacactgagccatgattggatc
actgcactccagcctgggtagacagagcaagacctgtctcaacaaaaatgaatgaaagagaaagaaagaaa
gagtgagaggagaggagatgaggggaggggagggtagcagggaggggggaggaaggaaggaaggaaggaag
gaaaaaagatgaaaaaagaaatagcaacatgaaacagaggcagaaagactttacgtaaatgctcatcat
gtggttgtcaagtttgacccccaaaacccaatttattgaccaaggttattctttgactgaggcaagggggtcc
gctctcctgggcttgggctttagaaagctcatctctggcctttctgagatccatccctttctttttatttt
tcttgacacggagtcttgctctgtcactcaggtggagtgcagtggcatgatctcgactcactgtaacctct
gcctcccgggttcaagcgattctcctgcctcagcctcctgagataacaggcgctcgccaccacatctggcta
atttttgtatttttagtaaaagactgggtttcatcatgttggccaggttgggttcgaactcctgacctgaggt
gagctgcccaccttggcctcccaagtgcctgggattacaggcatgagccactgcgcccagctcagatccatc
cctttctaaggggcaaacagtcctggtgcaaaggggcatgccaccagagttatgagtacctgggactcca
gaattccttgctggtggcctccacatgcacttccagggcctgcttgggctcttctatgggtctgtcctga
gtgtgtagaaccactgatgtgagtacctgggcttgagccgtggcctggagatcctgtgactgtagcatg
gagggggttgtgcagctgaatgtctgyatgcaggtgggtgggagttctggaatatgatggagctggaggtgg
gaagagaagtaggcttggggcagctctctcatgccacctcattctggccaaaactcaggtcaaactgtgaag
agtctaaatgtgaatctgcccctcaaggtggctacaaaggtatctttgtcaaggtaggagacctgtggcct
ccagctgcacttccagggcctgcttgggctcttctacgggtctgtcctgagtcttctatgzaatctgtcct
tcagggcagattcataatttagactcttcacagtttgacctgagttttggccagaataaggtgacatttagtt
tgttggcttgatggatgacttaaatatttagacatatggtgtgtaggcctgcattcctactcttgccctttt
ttttgccccctcagtggttttgggtagttttgtctccctacagccaaaggcaaacagakaagttggaggtctg
gagtggtacataattttacacgactgcaattctctggctgcacttcacaaatgtatacaaaactaaatacaa
gtcctgtgtttttatcacaggagggtgatcaatataatgaaattaaaagggggtggtccatattgttctg
tgtttttgtttgtttgtttctttzzztzzztzzztgtttttgtggcctccttccctcctcaatttatgaagag
aagcagtaagatgttctctcgggtcctctgagggacctggggagctcaggctgggaatctccaaggcagta
ggtcgctatcaaaaaatcaaagtcagggttgtgggggggaaaaacaaaagcagccattaccagaggactgt
ccgcttccccctcaccacagcctaggcctttgaaaggaaacaaaagacaaaatgattggcgtcctga
gggagattcagcctagagctctctctcccccaatccctccctcggctgaggaaactaacaaggaaaaaaa
aattgcggaaagcaggatttagaggaagcaaatccactgggtgccttggctgcgggaacgtggactagag
agtctgcggcgagccccagcgccttcccgcgcgtcttaggcggcgggcccgggcgggggaagggg
acgcagaccgaggacctaaagacacctgctgtacctccazzzczzccccacccacccacctcccccaac

tccctagatgtgtcgtgggcggtgaacgtcgcccggttaagggcgggcccggtccacgtgctttctgc
 tgagtgaactacataaacagagggcggaacggggcgaggaggagAGAGCACAGGCTTTGACCGAT
 AGTAACCTCTGCGCTCGGTGCAGCCGAATCTATAAAAGGAACTAGTCCCGGCAAAACCCCGTAATTGCGAG
 CGAGAGTGAGTGGGGCCGGGACCCGAGAGCCGAGCCGACCTTCTCTCCCGGGCTGCGGCAGGGCAGGGCG
 GGGAGCTCCGCGCACCAACAGAGCCGGTTCTCAGGGCGCTTTGCTCCTTGTTTTTCCCGGTTCTGTTTTT
 TCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGAGAAAGAGACGCAMACACAAAAGTGAAAAACAGgt aaga
 ggctctccagtgaacttacttgggcttattgtttgtttcgaggccaaggaggcttcgggaagtgtcggtt
 tcggggactttgatccggagccccacatccccaccacttgcaactcagatgggacggaggcggtgttaaat
 ggggagacgatgtcctagtacgagctctggtgacccaggactctgcgctgctgcgcttgzgggcttgcccg
 acggtggagaccggggagcatctctgggctggagaccggggcgagtagcccggttcagaggggtcgggg
 gttcccgzgcgtgctgagggcgctgctgcgggtggggagagctgcaggtccggcacgagzcgctgcttt
 gttcggaggggccctgagctggcztagzaaaccttctggttgaggtcgccagtagctacacggagacaaatg
 ccagcactgagttctcactcggttcttaagaagctggtctgttctgacctgggaattggctatatgtcccc
 gggactggagcggcacagttccggactgtgaatccgggaactcgagttggaggtgtcccaaaggctcggtgg
 tgctattgtcactagaggccttgggtctttgtzttgacctgaggggtaggaggtcctgcctacagtctcc
 gtgcgctcagctgagctggtgtccctggcgagagcgaggacgagttttgtttctttttcttttttt
 ttctttcttttaagtctcggtctgtcgccaggctggagtgcaatggaacgzatctccgctcactgcaacct
 ccgctcccggttcaagcgattctcctgcctcagcctcctgagtagctgggattacaggcgctcaccaca
 tccagctaattttttagttagagacggggttcaacttggtggccaggctggtctcgaacctcga
 cctcaggtgatccacgzgcctcgccctcccaagtgtgggattataggcgtgagccaccgcgcgggcg
 agttttgtttcttttaaaacaagacttaggagagcctgcggagaccggagggtggggtgccaatcctccc
 tctcccagttccctgcagccccatctccagaccgttgctgctggtctctcggggagcttctgctgggc
 gcagatggggaagctgggcccagggtggtgcggtggaatgaccgggagtaacccggcgggcgcgagaact
 cggagctccgcccgggggtgggtgggtctgcggtgaggggtgggggtgctgggcgcgcggggtcggtgg
 zccccggagactggccgcagzgcctcctggccggaggacctaggaatcgccgggtctactagggtctct
 tgcgcggttcgactgtgaatccgggtgaagaccgggtggttgagacggggaggaactatgaggtgagggc
 gaaagccggttttgtttttttttttttgttttttgggtttttttttttagctgtttgccaactcccag
 gccattggtaaagcaggaaggttcttggggcgggcgagcgggtgacagggttatgtgtaggtgctcttaggt
 atatcttttatcaaaaagaagcaaaataagattaaaaataaacaagaaaaaagtgtgtggtgactggc
 agtaattggcctgcttttgcagcactgataccattagcttttaaaatccgacttttcattgacacttcaaga
 agagaatgggttagtatatacacattcatctcatagtggacaaatttcatatttaaaaaaaccttctgggtac
 tgaaatcagcaagtcacttgccctccatggccgaatccctgcttcccacgaagagaacctcacaaaaatttc
 cccaagttaaagagtggaaattttcttgatttttttztcttttttttttaacggccgtagtttagaazccc
 agacttaattatgatcttcttttcaacaaaaacttaagtccttaagttttcatctccccttttatttcaa
 cctattcttctcatacctaccacaaaaataatggaggctttctgttgagaaactttccgtttctgttgagag
 tatcattctcttgagaaactttctcctaaatcagagaaagtatggaagcatggaaggtattcctgagtagaa
 cctctacagatattacaatatttttcaaatacaaagtttccattgtcagcctgtttcccaagtgttccaca
 aaccattaaataattccacaaaccattaaaataattaatgctagggaaatttttaggaaaacattgggtttacaa
 tcagaaggaccggggaagtgggtcttcagccttcacgatgactacaagccatttaagggacatagaattgtcta
 ctgttgtcagagcaatttaggagctgtgtatttgagcaccgcagatagttccagaatgacatatcagctgt
 aacctggacagctgtgatattgttctccccctgcagatgagcatttgaaatctcaacctcgtatttctacg
 agtgcaggcctataatggaccctgggcacatttttttttttttgagatgcagctctgcctctgtttcccagg
 ctggagtgcagtggcacgatattggctcactgcaacctccacctcctgggttcaagagattctcctgcctcag
 cctcctgagtagctgggactataggcgacgcaccatgctggctagtttttgattttcagtagagacag
 agtttcaccatgttggccaggatggtctcgatttctgacctcttgatccaccgcctcggcctcccaaagt
 gctgggattacaggcgtgaggcaccgcgcggccgacctggacacattttgacttagaacatatttccggttt
 gtgtgagacagtgcattagtgaggattggaaaagagttagcaggaattgattgtttcaaggattgggttcc
 ttctgctcaaggaagtccattgttaacataaaaaaatgaatgaaactgaagaagttcagtgacttagcttt
 ttattatcttctgtagtacttaocttttggagaggagttgggtgggatatttttccatttaaaattttttt
 ttaaagggtattctctcccgtgaagccgggatacttaagctatatatgtagtggctacaaattaagggtcttc
 actgttttcattttttagctgctagaataagtgaacattaccttagatagactcttctaattatgaagatata
 tagatgtctagaaaatatcaaaatgcatgtggtttttgcattttctaaaatacttttaaaacaaatactttt
 tctttttttttttttctgagatggagtcttgccttttgcctaggctggagcgcagtagaatgatcttggc
 tcaactgcaactccgcctctcaggttcaggtgatttctcctgcctcaacctcctgagtagctgggattacagg
 tgcgtgccaccaccccggttagttttgtgttttagtagggacagggttccaccatgttggccaggctgg
 tctcaagctcctgacctcaagtgatctgcccagcctcagcctcccaaagtgtgggattacaggcatgagcca
 ccacacctggcctcaaatacatttttttaagtatccagatatataaataaataataaccattatagtagttgtt

bioRxiv preprint doi: <https://doi.org/10.1101/151111>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

atgggtcatttactctagcatcaaagtgttaaaagatcattctgaacacttggtttgtttatgctgagagaagg
cctactccaaaaaatgcaaccatcttcgtatctgcatgtggatacaaccatgaatggccaaagtatttgcag
tggtgaatagacacttatatagcactgtgtggcaagtactgtttgaaatgttttccagtggtttattcattg
tattttattttgagataggggtcttgcctgtgtgcacaggggtggagtgacagctgcacagacaagggtcactgca
acctcagcctcctgcgtcacgtgatcctcatacatcagcctcttaaggagctggggccacaggcacgcacc
actgctcctgggtcaaatTTTTTacaatTTTTTgtaaagacaagggtctcactatcttgctcagactgggtcttga
actcctggggtcaagtgatcctcccacatcagcctcccaaagtgtggtgattacaggcataagccactgtgc
agggtcatcaaagtctaattgaattttcacacaacaaaccatttattgtccctagtttacaagattaagtaaat
gagaagctaatttttctctggctatataccttgcaaggagcagagctaagacttgaaccagccagaggttct
ttaactccagcactaacatttcagctgctgcaaccagggagcttttcaaggatgatcaccacattctctaca
ttcatctgctataatccttatcagaatctacagcctgtatcatattttccttggtgctgtgagtggtgcagc
caaattctctttaacttgaaaccttggttgctgtagggattgcaacatcctggaaagaatagaataaaattta
ctcaactcaattttttacttggttcataatgaaaactatactattgcttcagtcagatgtttgcgaatagct
gtgtgatctcaaatgttttctatgtgatctatagzaaaatggaatgatagagtattaggctgtaagggcc
taagazaacaaaggaaaaagagaagtgaactgttagtttagttgttaaaaccttaactttggtgaattgtaaaa
atttgtttataatacaatatgattcttgcttgctcctgtccttgatgaagttgtggaccttttgaaataagcta
tttctctgttactgctgttactgttttagaaatcaaatttagtttttcttaagatatacgtatttttggaag
ataaacacagtttcaaagtctgccttggttggtggtgactgggtcattgttgattcccagcactttggg
aggccaaagcaggaggatcacttgagggtcaggzagttzcaaagaczagzcctggcaaatatggtgaaaccc
cgtctctactaacaatacaaaaattagctgggctggtggtggtggtgcttgtaatcccagctactgggattgg
gaggctgaagtagaagaattgcttgaacctgggaggcggagggtgcactgagtcgagatcgtgccactttac
tccaacctgggagacagagtgcagactccgtcttgaaaaaaaatgtctgccttgtaaaagtgaataggatga
gaaagtgcctttcttatttaattggtgtaattgaattagaaataaactctttgaagacacctcttggtaaaaat
agttacatttactgttgatttatggtatgttggtatgtttttagttttccgtgtaataactcagttcatt
ctcatgagtgaatatagggtgcttttattgtctttatagatgggaaactgaggtataggcaggttaggtacatt
attatggagttcgttaagtgtggagctgaagtcgcatcccagacagtttggtctccgtgagtttaccatct
catggttaagactttgtcagactatcaaagttttgacaaatgaaatattagcaaaaggccaaaagggttct
ctattttcatttgagtatcttcacctgaaaatagttgctgaataagtagcctgcatagaagggtacatttt
agaaatacttgaggccagagaatgaaaagcttacataaaaattgattttccggtggggccttcagttactctcc
attctacgaagaccacaaatagcattcaggcaaaagagcatttattccaacaatggaggagcactggatttg
ttcctaaaaacaaaataaagtttgaaatcctgtctttcccatgttgaaaacaaagtgggtacaaaaccttta
gcttttgcaaacctcctttaagacccgatttaaatgcytccctcctcatgaagctcttctggatccactcyt
tcccatcactaagttgaaagtaagatccccctctcttacttccattagacttggattacagcactctttgt
atcatgtatttaattctgttttttaattacagtttaacatttatttgtcttctcttgagtgtatgcttctct
agaggaagggtctttgattcattctccccctggccttaattcatcccacttaatatggaaaaaatttaataaat
gctgacttgaataagtccaacaaggagaatgggaagctcatgtttgcttctctctctctctctctctctctct
agataacagggttaatcacagaaaagcattagaaatagagttatatgagaaacaactgtagtttaaggctagg
ttatgttagactgagaaatttttagtgcatacttaagttatttaggccagggttactttttagtagaacaacat
ttcagtttgcgtcagtttcatctccgtttctgzaggcagctgtgatttaagaaaatgctctagctctgtggca
ttccatattcaagtactttgagttgtatatttaattttttagttaataagagtgcactgactcactaagta
atttagagatttaaacacttttttaaaaaacagtaactctcatatgcattggatctattctctataaagctct
tttcttgggggggtgtttgttttaaaattccccggtgttttctctgccaatccaacttccaagaagcatttg
aagtcaaaaacattttatctggttagtcttaaaagtcagatattttgtgatagctgggtatttagtttatgata
tttcccaggaagaacttttttagtagttgaaccatttatgaaagacttcttgaaagctaccttagagagttga
tttagttcttctctaaataagtaaatagaatattagttataggacatcttgaggtatagatgcaaatattggt
gaaaaagaacatggatatacagagtcaaattaatgtagattggaattctggttattactaatggatatctgac
attaggcaagttgctgatcactctttgcctcagtttcatcatctgtaaaatagggtatttgtgtttgtgtata
atgtgaaccgtataataatgcttggcctatagtgaaatttattcatacagagtgttttcagtgattttaa
agcttgcttttaggcggggcgcgatgggtctctgcctataattccagcactttgggaggccaaagggtgggcggat
catgagggtcaggagttcgagaccagcctgaccagcatggtgaaaccccgctctctactaaaaatacaaaaatt
agctgggctggtggtgcaagcctgtaatcccagctactcaggagggttaggcagaagaatcgcttgaaccc
aggaggcagagattgcagtgagccgagatggtgccactgcacagagcgagactccatctcaaaaaacaaaa
caaaacaaaaacaaaatcttgctttatagtttacttccacatcaaattgtctttatcccatgttacttgc
tgatatcccagacatgaaaagaaaaaagatgataacaatgacagttattaaattagggttccactcttattc
tagataccaattcatattactattcagacttggaaacattaaattttagttaaactttttcaaatatgca
tataattgtcagtggttactataattttggggaagagattgttgacttctttgaagaaagatacggattttct
cttcagaaraaatacacatggggtcatataatccaaattttatgtgtaattacagggtgttcatgaatgccc

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted November 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

acaaatccattaagccatgtgaccttggacaagtcattttacttttctgttttttaggttgttggctctgttaa
aatgatactacttgacttttaagagcccttcaagctcttatgtcctctaaccacaggtctgtattcagaag
aaggggtggtcctttaattagagccatctagagatctgaggaacatgctgggcatttagtgtaacataccatg
tggattttgagaggttaaagaaaaataaccaggggaatgcctcagagcattcctgatcagatcgatgacagaa
gaaaggaatgagaggggagagaggaagctgttgaaatttctatattacctgctttgagtgaatgaagatttg
aatcatagaaccagaaggggttctcatctgaaatgcaaaggaaggaggaggttgggttaattcaataagtttc
agttgagtaaacatgatttagtgagatactgttcttgcctctgactcaccatttggaaaatctctctaaaat
aaaattggactctccatctcggacatcattttgggtgtaggttttgccttttttttggagatggagtcctcgt
tatgttgcacaggttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag
ttctcctgcctcaacctcctgagtagctgggactacaggcgtatgccaczcagtgccggctaatttttgat
ttttttaatagagacgggttttactatgggtggctaggctgggtctttaactcctgaccttgtgatctgccca
ccttggcttcccagagtgctgggattacagatgtgagccacagtgcccgccctaagttttacttcttataat
ggactcctgttaagccaataggtgatgaaaggaaaccataaccaactctcaggtcattcatccttcaaga
atagcatgctagtaccatcctaggaggagaattggactatacctcatgaggatagtttgaagtatctcaga
agacctcactggggtaggtcggttaagacacaaagctttctaaagcactgtaccaaatttgttgtttgaga
gatcatacaaattagaagtggaaagaagaaggagtaaaaggaagaagaggtttctggccaggagcagggag
ggggaaggagctgctaggaagatgtttgggtgtcatatccctgttcacccttgccttgcaaaattcttgtag
gatgccaggttgggagtaattgtttttcacaagagtcacacacgcttgttttctgaagaagcaagctctt
ggaggggtggtggcttgaatctgttggatctggtatattaggtgatacactttgacataaaggzcaacactga
tgcaagcagcagctttccttggaaaggcagggagaaagtgaaggccagactgatgagcttacactgacctg
cagaccttctccattcccaggtatgtttgggtggcagagttaccttagttggggtaggtgtgtctggt
actgtgagagagaaggaagaagaagatgatattaacaacaacaatacatatttatatttgaatttaaat
gcactaatcacctatagtgtctattaacaaatttatttgttcagcaaatgtttgttaaacaccaatgtact
gtggaaagtactaggtgctggacagaggtcaaaagactttttaagaatctgccaccattaatgatctcttct
tgcttggcattcaaggctctttgaaataagactgtgacccactttgatagttttgtcctggattataagaca
catgctcgaagggaactatagctggttttctcaccagactgattaacatatagtatgggttgggtacctgttaa
atgagctctctctacaggttttctctacttttaagaaccttctcctggcattgtttgggtcctcattg
ttctggaatctcatgtccatctcatgtatttgccttgggttcacacttgatctctctgctcatattcctcacct
actgaaattttacccatcaaccagacccgtgggtgatggaacacagcatgggctagttctcttataatgat
ctcatttaattttcactggaactctgagataggttagcatttcagccactcaagttgactcaaatttttgta
atcatcaatataattttaaaataactttatatttgacctgtaattggaaaaccaatagattatcataaatg
aaaggtaattttaaaataaattaggatgaagacaaattatttttctcacagtctatgtataagataaactat
tgggtcccaaaggccccagctgacaatgagacttctcttactttgttgaaaaggggaattagcaagcattaaa
gaggtgtcaaaaagaagactaaacaaaagcttactccttttttttttgagacgagtccttgctctgttgccc
aggctggagtgagtgccaccacctcgggtcactgcaacctcttctcctgggttcaagcgattctcctgcc
tcagcctccctagtagctgggattacaggtgcatgccaccacacccggctaatttttgtatttttagcagag
atgggggttcgtcatgttggccaggtggtctccaactcctgacctcaggtgatccaccacctcggcctccc
aaagtgtcagaatcacaggtgtgagccgccgacccggcgcttactccttaatgtactaagaatgttatat
ataggctgaagaagtgtgaaaagaaccatattttctcatgatgtggttcaatgtttaatactgtgcttgtt
catctcctaaaatcctctgaatatcacttaaatcatcctgtgtaactctccacatttggggaaatactga
gcttgccctattattatattagccccatatttcagatgatgcacctgagccgaggagaagttaaataactgt
tcatgggtccatatttggctaattggcagagccagggattcaaaactctgtctctctgactcccaggtttgtg
cttttcccacttgggtgaatttctcatgtctacctcctccataaacaccttctcctagaacttttaaggaatgct
tcccctgttctctcatagcatttttaagttaagttgccaaggtgtcccagtttgcacccctaccagcaatgtg
ggagaaaatagcaacatatttttgatgttgggtctagtgttacgggtttctctgctcttgggagtgatat
tccatggctactgaatacccaagtcaccaacagttttcttagactcagaaggtcatccacttcagcttcttc
atcaaaaagacatattttctgggtgggcaggtgggtcacacctgtaatcacagcactttgggaggtggagg
caggagaattgcttgaacccaggaggtgaagttgcagtgagccagatcgtgccactgccctccagcctgg
gtgacagagcgaaactctgtgtcaagaaacaaaaaaaggacttgttttctgttccattaccacagtggtta
gaatggcggtgctaaatttattctccagctgccattaaactgcaaattaaaatcttagtctcttgccctcttaa
tccaggtctcttcatactataccagaatttaggataactattacagtgccctttataggagagaaagaagaa
attgtgtctgtagatgtctgttcttctcagcttaaaatggacactgaaatgttaaatattggactggcctca
tttattttctcctgtctgttgggtccaatttgaatcttaaggcgtctttcaactggaatttttgtttctctca
actaaaatgttctttgttaagtttgaatcagaacaaaatcctgaatgttgagggtttcctaaaggctgttt
ctttatgcaaaagcctgaaacccgatgttgatgttgggtgcttaaaattaactgtgaatcaaggcaggggtt
ttatttttatttttttttacttttaattgattgtgttaattatagtgaaaaccttgaggtcacgagaaagaaa
gcctttgggtcaagtatgttttattaagttgtcagctctgttgcaggatttgcaaattagtggaatttagtgc

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

catttttcagtttacaattccagtcacatttcacatgatcagagcatggctttcttctctgtggagcaaata
gagggctgtctgacacttgggtccagtggttccattaagcagagtggatatgtccctggagctctgcagaga
agggcatggcactctgaccccagatggcactccggttttgggacattgtccaattctagttcatagcatatgt
gaccaacaccagctctcacctgatgtaaacacttagcgcggttgttgcctgggggattggattgtgtgaattt
ttcaaaactacagttgacagaaggaggctacaaaaatgaaacccaataattccatttttgggaattattcc
cacttttgttccatttttcccactttgttctttggcacacagaatgtttgatttgtgaaaatcttaataaca
gtagttttttctataaggaacactcagaatcttgataatattggaataatacagatcctttttaggatcct
ctcagacctcatataatagagttcatgttagtcaatatttaaagaaaaacacccttaagtttttgttttccag
aatcacaaagtaagtggatttaaacttgtgatcttattccccctttcttcttcttaatttagtgaggcagccagc
gagaggggttgttttgggtattctaaagaaggagtttgccttgtaagttttggaggggaagacttagactctg
tgtctctgtgcttgccttggaaactttgattaaattgtcactaaccgagttagctggccctcgccgggtgca
gaaatagaagtgtcttgcacacatgacatatgactgtctcaagagctggctgggtgaaaggacgttctggaga
aggctgccgatactgtatgaactagaazctggacaagagcctggagattggataactcagtttggcgcaagt
aaagggaataaaaagtgttaagggtggcaaaattgtatccagggtgtttataggctccctgagttcctgacttga
gcctatctatgggttttagagttcaaggctctttaccagtgctgacaatcttatactctaggttgaacctccg
gggaagggtgcccttgccttgatggcatgtttaccaggggttctagagcctcaatcacagattctctctagctc
acatgaagttaatgaaaatgaatgtgcttccctacaaattagagaggctttgaggaaaaatcagattaaatg
cactcctgcttgaacttatgtttcttagaacacagctggaaattttgtcacacaaacctttactttcagtga
catttcttgactgggttgttactgttagtgaatctgctttaactatcttttcttatcgctgaggttttacttc
cattctacatgtgattgtggagcgctgcgtcattgtgggttcagtgtagtggagagtaggaagatgggtgaga
cacagtagcttgttgcacattgcttaatttatcagggatcactgatgagttagtacactagagaagattgta
ggtagagctgaaaagatggaggaattataaggctcagatttctctcttttttttttttttttaagatggagtt
tcaactctgttgtccaggctggaattcaatggcatgatctcggtcactgcaacctctgcctcccggttca
agttagagcttgatgggtctcacttgatgggttcttgggtcagcctcctgagtagctgagattacaggcacc
ctaccatgccagctacttttttgtatttttagtgaggatgggggttttatcatgttgaccaggctgggtctcg
aacttctgacctcagggtgatcacctgcctcagcctcccaagtgacagggttacaggcatgagccactgcgt
ctzggccaaggctcagatttctaataagagatzttctaattggacatagaggctggaggaaatgggatzggaca
zggazaaactgagtzczaggtgccaaazzaacttgtagggggccgggtgcggtggctcazzcgzcczctgt
aatcccagcactttgggaggtcagggcgggtggatcacagggtcaggagatcgagaccatcctggctaacac
ggcgaaaccccgctctctactaaaaatacaaaaaatttagccgggcgtgttggcaggcgggtgtcggtagtccc
agctactcgggaggtcagggcaggagaatggcgtgaagccgggaggcggagcttgcatgagccgagatcgc
gccactgcactccagcctgggagacagagccagactccatctaaataaataaataagataaataaataaata
aaataaaaacttgtagggaggtggcagtgctatggaggataggtgcaacctctgtgagaatgtagagaaa
atagtataagttagtggtgaggaccccccaagaggggttttatagtaaaacaatggtcagaagtggcaacag
gataccgtataatgctttcacctctaccaatgcactgggtactggagagcgctccagtttgctctggaaagg
ccctttctgtggacaaagaatacagaaaagagattcctttaataaaccaccactctgtgtcccacccttga
tgaatacttcaactgtgaaattgccagaattaatcatggtaatagctactgtacacttactttgttccaggaa
ctggataaatgttttacatacattatcagttctatttttggagaagatacaggggctcagagtccctaggg
ttccagagctgggtgagtagcagagtcaggattcaaaccagctttctctgactctaaaacctcctttcttcc
tgctgaaacaattaactcaagacaacaaaggagtttaaggatttggggagttttctgcatggtagaatagacc
caaaggaaaagaaaagacagtgactaagatttgggttgcctgctgcccaccaaattgctttgagcactttc
ataaatataaatccttcagggttgggagaagggttgaacatctgaagacactgattcttcagagatgtaatcca
aacaagtgatctttgggtgatattggtcactaaaccatttatcccaaaattctcttggaaaaccgtcctataa
cagcaggggaacattatccagccaagtttttctgcaataaagggttgcctgatagaggcttgctgcttgtgt
ttctzgtzagcztzcagggtgtttatgaattcactaatcccttcccttcagatcccttttattctgggtgtta
tgattgtgactgzaaaaaaattgatttttttctatgacatagaatgttgaaagggtgattttcttttctaga
ggaaagattcttttttctatgtgctacataccccccgaccagggaaaaggcaaatagtgggtattgtttgct
gaagtcttcccttgaagggttgccttgggtgtttgcttagtggaatcagcaggggaagagaggctatctctzaa
cattttgttagazgtttcttctgtagttctatagtgatgaacaaggacttggggtagaggacagatctgct
ttcagaaatcctggctcttgtgaggtttagaagccctgagaccatttagctgggtggcaacgggaatgttgag
gggtgataaataggatctttgggtgtccaagtatcagtgacatgatatagatggagttaaaccttttaggatct
ccttattttatttgtttgtttatttttgagacagggtcttgcactgtagcccagggttgaggtatgggtggcatg
atcataactcactgcagcctcaaactcctgggtcgaagcgtcctgctgcctcagcctcccaagggtgttagg
attacaggcatgagccgccacaccgggtcaggatctcctgtaaaattatattgttgacaacatgaagaatta
tgcttctcaaaagctagttatagatttgtacaatattcatagatttcttgtttcagtttttcaaaattcata
gcccttatttttggaaaattagctatttagcaataatttgccttaggaaattggatgtgtattcaagtgaagaa
ggaagtacagttacctattatcttattgttaactaacaatcaagtgaagtgtgatttgggtactttaaaaa

ctgcacccaagttacagattattggaattaataaaattcactggatctatatatTTTTTaaacggacagtgtg
atagcagaacctcttatagaatzgatagaattcctctggaatgattggataacttcatttcaccttgactt
ttaccttgaggatttcttacccttttggcttctcaaatTTgactattaaatgttgccctttaaataagg
aacacagtttcaggggggagtagcagcccatgacccttctgcaaggccccctaactcaaggtagtttccctg
gaactgtggtttatggaatgtttcaggagtgtagggagggtataatttaaggctgtcctagcaaggataccct
taaggatagagggccagtagcatctggaggccagaaaagttazaactgaggcagtcagattagcttcazgg
ctcaattaagctgatgggtcagcctgggagaaaattgcaggatgactctcaatatccccctccacccccacag
cagccacgatctgtctgtctttaatcatgggtgcagtgaacctgttcttccagggtgtcttggccttcagta
accttgttaggcttgtccctgaacgtggctaccgatccaaagacacatgatcagagagggaattagagaaca
gaccttttccaaagcaagcatgttctgttgggcttagaagtttcatgtcctaataattataggacctgtgca
tctctctggagatgaggcacatgagtcataatctgtgattcttggcttttgtgtcaacatctcatgaataggca
atcagagctttggcaccaatgtattttcagttcatatctgatgtagttaaattccacctcctgtttgtagtt
tactggcaagctgtttttgatataagacatctagaacactgtaaatatataacattttttatttgtctattat
acctcaattacgaaaaagacatctagaagcaacctcatcaagagagataactgaggccgggcatggtagctca
cacttgcaatcccattactttgggaggctgaggcaggtagatcacttgagggtcaagagtttgaaaccagcct
ggccaacatgttgaaacctgtctctattaaaaatacaaaaaagttagctgggcttgggtgggacactgt
aatcccagctactccggaggctgaggcaggagaatcacttgaacctgggaggcagaggttgcagtgaactga
gatcacaccactgcactccaacctgggcaccagagttagattacatctaaaaataaaaaaagtaataaaaa
aagagagataattgatagctgttgttggaaatttcaacttccatctcacttcttggttaacttttggagtttg
ttgaacaaagtggaaatacacgcacatacacacacacatactctcttgtttgtttaagggttaataaaata
gctgtcatataatcactgtttttgaaagaggagaattagttgctatctgtacatttttgggtatgtgaactat
ttggatagaactctgagaaatgcattcagaacaacaaacaaaatcataggagaaatagctaagtgggaaggg
gcatataagagttgttgaaaaagttatttcttgagaaaccagctctaagctaggcaagtcacttgctttgg
gggaggcctcagcttctctgtctataagattgcagcaggggtgtagtgggaatgagcttccaacattccaag
agattttatctactaatacagcagtcgaatggagcatgactttgtggaagcctctcctcttccaccagagg
ggccaatttctctgtcccagtgagatgttgacacttgtagtacctgcttggagacttccctcttctgga
cctgcccctggctcaggcatgagggtgactgtcaccctcgataggagccagcactaaagctcatgtgtg
gcagtgttcttgcgggaaggaaaaagaccagcccatattgttactgcacaagcaacagcttctggtag
ctgtacagatacatgcacttcttctcactgtgtttccatagacagatttagtgctgtagaagagtagag
ggcagtcacgggaaggagttcctgtttttcttggctatgccaaatggggaaaaatcctcctatctgtct
tttttagtgctcactctctctcccttttcttcttcttataattctcatctctcctctcctggaaatgtgc
atgtcaagttcaaaagggcacaatgttttgggtgaggaagaggtgggagaacacgtgccaggtgctaactagg
gtcatcatttcccccttcacagccagcttctgtgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
gtgtgtgtgtgtatttcttttggcagcatcactgaatctgtctgtgtgtgtgtgtgtgtgtgtgtgtgt
ggaaaagtaaaagtaattttataatcccagctgtcatttaagccaccccztgttgggtagcatatgggtcca
ctctctcagttcattgtcctaaagatgcttcacagaaaggaataacttccaccccgtaactctctgtcccc
ttactctgctttatttttcttctgcaatcctaccaccaccaccactgtttgaacaaccactattatttgt
ctgtttcccatccctggtagaataggagccccatgaatgaaggaactttgcttctgttgttccaccactgaat
ctctaagggtatggaacacacctggcatgtgataggcactcgataaatatttgttgtggctcatgggcacctt
gcagagtttaaggctgcagttgtttgtggaatttataagtggtaatgaatatttatctactattcctcttcca
aggcgtcacacaataatcaggctttacactatccagttcttaggtcttccaagttatgacttgtgaggtat
gttaattatgataatagaaggcagtttatttgggttcagatttattgatgtgtaatttaccacagtaagactt
cccccttacaaaagtatgatgagttttgacaaatggatcacatgtgtatctaccactgccatgctcctttt
cagtcgtctgtccccctccacccatgaccactggtcaccactgcagtgatttctgtcccccttcatttcacctt
ttccagaatgtcatataaatggaatcatgcagtatgtagttttttgtgtctggcttatttttcttagcatta
ggcttttgggattcatccaggttgtcgcagtgaacagtagcttattcctttttatggctgagtaagtgtccc
agtttttatttatataatttttattatgaggaggtgtctcactctgtcaccaggtggagtgcggtagcgcgat
ctcagctcactgcaacctccgcctcccaggttcaagcaattctcctgcctcctgagttagctgggattacagg
caccaccgcccacgcccactaatttttatatttttagtagagatgggggttccacatgttggccaggtga
tctcaaaactcttgacctcaggtgatccgcccacctctggctcccaaagtgttaggattacaggcatgagcca
ctgtgcccagccccagttttattttaccagttgatggctcttttcgacaactaattgtttccagtttttg
gctattctgtataaggcttctataaatattcacaaatacctaggatgggatgactgggtcatataatagtac
tgtataaccttagcagaaactgtcaaaactattttccaaagtggctcttccattttacaattccacagtgat
tgagtcccagtgctccatacacatgctagcacttttaatatttaatttagtgggtatgtaatgatctca
ttgtgggttttaatttgcatttctctgcagctaattgatgagtggttctgcttatttgggaaggttttaattta
gcagtcgtgttattctgtagataatttaacttcaaaatatcagtggcatttgcagttaaaatttctttaa
aaaattggccaaagggttccagcagtcacttctgccatgccc aaactgtatgaaacaaggctgaggtgtgga

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

gattgtcacattttggcaaggagtgatccacttgggtgactgatgagaccagagagcgtagcctcgggct
tgagggtgaggacgggagggaagtcgactgcatggccctgctggccttgggaggctgccagtccttagcta
aagctggcagttatgggaaacagacttagattctattacgttttccaggatgtcccaggagtcacctgggaa
gctcagcagtcctttgtgactttcaagcatatggtagaagctgctgaacacagagctccctctttggggata
atgtgccaaatcatttaacaggttgagaaatgagttaccacaggtccaggagtgctgccacccttgaat
tctgacaccctatttctcctatccgtctcttaattaattaagcagacatccccaagtgcttacgacaagcca
ggacccttttgcatactaaggaaaacagggatgaaggaaaacagaaatggctctgctctgactcagaaggta
gaaatcctctttccagccaagtcctcctagggagcacgttaggaagggtctgaaccacagtgctcagttgca
ggggaggatatacaggaaaggacattgaagaagtggagacctaggtttgagacctaggcattagccaggctag
cagtgcttgaaaaagtgtcttaggacaagagaactcaccagtgaggtccagtggtaggagagcgtgcagca
tattctgagcctgtatacacatctccaggcatttgcttagcaggtggggagtggaagagtaggctggag
tcacagaaggaggccaggttagaccttggtagcactggactctatgttcaggtgctgaggagctggcaaaa
ggttttaagtcggggagaggcatgttcagatatttgggtctagctgagtaactttgggtgctctgtgacaaat
ggttgggagaccagtgaggtggcagttgcggtcatctaggagcaggatcagagtggtcctattgactgggatg
actgtgaagtgggatcctttccagccagtaactggaaatgtgtatgagggcagaagtgagtgactgcattt
gaaacattgagaaatctagtacatagtagtctcttttatatcttttttttttttttttttttgattttgggtt
tgtttgttctactaacttggaaaactgatgtggaatgtccctttgggttcagttacctgagcagaaggggccc
gggcattgccaactctcctcttaggacagaattgtctccagttatgatcattgtgttctgagttgggggag
caaattgtgcaggaggccaggtcagtgccaaggtgggtgggaggaattggagcaggaagcttgccaaagtg
gccagcaaagccacggtagaactttctactgtggctctatgctacttcttagcaaccttctccatgtgctt
cctggagagtccttggagtcagaacctttttcttgaaaccagacactttacttccaagaaaatgctgtcca
agaaaactcatccttcccttcttctcatgaacgttgtgttagaggtgtgtcttctcttcttcttgagctttcc
actcagggttttaggggaggtgatattctatatttgggtttgggtctgggtactgcaacactaggctattaag
atctcatccttactgctttgcccctcctatctttccagaaaccacaatggatttgctagaaataatggaac
gtcctgtttggacaggatataaccatttctcagctagaggatattgttggaatgaagaaagataaatgggga
gaagggaaactcacattgtcttggcacttaaatgaagcattagcactgtgtgttgggaaattatttatattctc
gttgaatccacagtagaacaacagttgaacaccataccaaggttaagttgtcatccttattttaccatgagga
aattgatgcttagagagcataaagccttggccagggggcacatagttgggaagccgggggtaattcatgcctg
ggctcttctgatagttttccttttttaattgtccctcctcattgttaccttggggatttcaagagattca
tgtagcttctaaatcaacgaactgattcctggagagcagcttctgtatgagaaaaatctagctaattattta
tttcagtgctctctggaatgcaagctctgtcctgagccacttagaaaacaatttgggatgacaagcatgtgtc
tcacaatgctgctctgggttggcagtgctgtgctgccagttgtcatcttgaacaaactgatgcagtgctgggt
ttaactcttctctttttggagtaagaaacttggaggcctgtgtccttctagaagtttgcctgagcaaatgg
taagggaaagaaataggtcctaaggcttgactatttcagagaatttcttgatttattggactgtcaatgaat
gaattggaatacatagtggttaggtgtcttttctctcagacactgcaatttccctccaatctcttgactttt
ctagaagttttaatccaagtccttgttgggtgggtzagataaaaagggtattgttctactagagactgacctg
gcatggagatctcatttggactcacagatttctagcttagcgttgggtttgtatccatacctcgctactgc
attcttagttccttctgctccttgttccctcatgccagtgctccaccctacccttggccctactcctctaga
ggccacagtgattcactgagccatttcataagcacagctaggagagttcatggctaccaagtgccagcaggg
ccgaattttcactgtgtgtcctccctccatttttcatcttctgcccctccccagcttttaactttaatat
aactacttgggactattccagcattaaataagggttaactgctggatgggtgggtgggatcacagaaatgtag
tatcccttgttccagagaagaccttcttgcctagcatggcaaacagtcctccaaggaggcacctgtgacac
ccaacggagtagggggggcggtgtgttcaggtgcaggtggaacaaggccagaagtgatgtgctgacca
tgggagcttgtttgtcggtttcacagttgatgccctgagcctgccatagcagacttgtttctccatgggatg
ctgttttctttccagagacacagcgctagggttgtcctcattacctgagagccaggtgtcggtagcattttc
ttgggtgtttactcacactcatctaaggcacgttgtgtgttttccagattaggaaactgctttattgatgggtgc
tttttttttttttttttttgagacagagtcctcgctctgtcgccatgctggagtgtagtggcacaatcttgggtc
actgcacctccgctgccaggttcagcgattctcctgctcagcctcccaagtagctgggactacaggtgcc
tgccaccatgccagctaattttttagtatttttagtagagacgggggttccaccgtattgggtaggatgggtctc
gatttcttgacctcgatcgccctgcctcggcctcccaaagtgctgggattataggcttgagccaccacgc
ctggccgatgggtgctttttatcatttgaaggactcagttgtataaaccactgaaaattagtatgtaaggaag
ttcagggaatagtatagtcactccaggttgaggcaaaatttacaatgctgctgactttgtatgtaaggg
gaggcattttcttagaazaagzagaggttaggtctctgggattccagtatgccatttccatcctcagtggtttt
tggccacctgagagaggtctattttcagaaatgcattcttccatccagatgataacatctatagaactaaa
atgattaggaccataacacgtagctcctagcctgctgtcggaacacctcccgagtcctcttcttgggtgaa
cccagaggctgggagctgggtgactcatgatccattgagaagcagtcagtgatgcagagctgtgtgttggaggt
ctcagctgagaggggtggattagcagtcctcattgggtgtatgggtttgcagcaataactgatgggtgtttcc

aaactgagtcctcactctattgcccaggctggagtgagtggtatgatctcagctcgccgcaacctctgcctc
ccgggatcaagcaattgtcatgcctcagcctccgggtagctgggattatgggcacacagcaccacgcctgg
ctaattttttgtatttttagtagagacatggttttagcatggtggccaggctggcttgaactcctgacctca
agtgatccaccacctcagcctccaatctgctgggattacaggcatgagccactgcaccagccagagtac
cactatttgggcattctttaatgaaaaagaatgaactatccaaaaattaaaactcctcatttatgagctttt
agagaattttacagagtagatggaaactctctgcatcctttcccccacttctagtttcacctgacacatttct
tccctgtccttactcctgggcccggcagcagtggtcatgattccaatcccagcttggccaccatctgcctcag
tggcctaggaaaaactcctttctccagagcttttagttttctcttctacggaatgaagaaagttaaaacaaata
gacattttattgtttcatttggataaatatctattaagcatctattacttgtgggtatgggttagctgggtatat
agtggtagagcagctgggcatgagtactgctttcgttagagctttacagttcagtgaggccagcagatgtgaaa
catatcatcacacaaataaaaaataaactatcaactgtgtaggattatgaaggaaaaaatccggcaaaact
atgggtactgggtgttagatactagcaggtgtgggttagggatttcatttagattgacaggttgtcacattaaag
ctgagagccctgaagttcaagcaatgggttagccaggcaaaagatcagaggcttagagatagggaaatccattc
caggcagagagactgggggtgcctgtcccctaggctcagggaacagaagaaagccagtggtggtggagt
aataagactggcgggggatgagttggtagtagacatgaccagatmatttagggczcaattctcztggggza
aggagaattztaatttaataattttattttattttattttattttattttattttattttattttttcaa
gacggagtctagttctgtcgcccaggctggagtgagtgagcaatctcggtcactgcaacttttgctcc
tggtttcaagcgattctzcctgcctcagzcctcctagtagctgggattacagacgcccaccaccatgcccag
ctaaaztttttgtatttttatagagatgcgggtttcaccatattggccaggctggctgaaactcctgacctg
tgatcctcctacctcggcttcccaaagtgtgggattacaggcgtgaccacagtgccccctgagaatttaa
ttttattttatgtgcaagaggattccctgaggtagtcaggccacattgtctggtgactcttgggtagaggg
aacttgaatgacaaaggcccaagaaagcaattgtaatcattacatatacatggaccattttatgctgttttc
ttctttcatttaacattatttagtggtgctgttcacatatttctaaatcatcttctgatttagaataatgat
ttctgatgtgtaggctgtgttttatagttttgaaagtaataactttgatattcattacttztcttgattctca
cagcaattctgzaggtgtatgcgttgcaatttctgtttcacagatgaagagagtattgttaataagttaat
gcccgggcatgggtggctcacacctataattccagcacttggggagaccaaggtgggaggatcacttgaggcca
ggaatttgagaccagcctgggtcaatgtgggtgacacccatctctactaaaaatacaaaaattagccaggctg
gtagcacttgctgtaatcccagctatttgggagggtgaggcaggagaatttgcctgaaacctggcagggtgga
agttgcagtgagccaagattgcaccactgtactcctgcctgggtgacagagcgagactctgtctcaaaaaat
aaaaagttgctaagaggagggtggtggtatcttttggctccaaatctactgtgggatgatgcctttgacattcc
tgatagctgtgcagtaatccattaacacagtttttataagttcaaaazccctgttgccaacatttagattgtt
ccatgtgtgctgttacaataaattactataaagattctatacatttaattcttttattttttgtattatt
tctgtaggccaaaatctgaggaacaggattactaggttgaagggaattggcccttgaagtgtctgatcagat
gtctttccagaggatccaaccaatttaaatagccaccatcaatgcatgagactttgtagttcagggaaggca
ggcctgggttttaaaaaatcatttccccctctctagcattttttctgatgtgatccttaagatttcactttagttt
tcccagggtctcattggcatgtatgctgttagggatgggtctaaaattaaattttttcttcacattcatatcatg
tcatcccagtgattatttaataaataatcacttgattaaatagtgattccttttctagttattttttgggaca
tttattaaaacctggataggtggctcatgcctgtattcccagcacttggggaggctgaggtggggggatgtg
cttgagactaggagttcaacaccagcctgggcagcatagcaagactccatctctataaaaaataaggaaatta
gtcaggcatgggtgggtacttgctggagtcagcagctattgggaaggctgaggtaggagaattgcttgagtcc
agggtggtcaaggctgcagtgagctatgaccatactactgtactccagcctgggcaacagagtgaactctgt
ctgaaaaaaaaaaaaaaaaaaaaaaaaaaaaazzzzzzaaaaaaaaaagatgtgtaggagcaattttggagttat
tcatttgggtcatttgatatgtagtttttagttttgggtgctgatagagccagaatgtaccctgaatttgatga
acattctgatatatgggggagctcattgtccccacttacctttttgcctctcagaatatcttttgatattt
ttatctgttttttccccattgaatgttattaccttatcaagctcaaaaaagtagcctatcgctattttaagt
tcagttgtgttaaattctataaattagcttgggaaatttggatattaaatgaactcatgaagaagcagagttt
agctctccttaattctcatcttccctttattttatctactacagttctgtgggttttcttttatgtaagaagca
catgztttggctaagttaatgcctagggttttttggttatgtgtccattctcactgtggatagttctctt
tccccacattatatttaatttaactgggttttcagagactaatagcaatgctattatttaggagaatttacct
tgggtctgattaaacttaccatacttgcaaatcatttgcagcttttttagttaactttgtgagttctcttaga
tttacgaccatgccagaaacagaaaggatattttcatctcttcccttctgatgtttattcttcttgtttcct
ttttttatccccattatattctcaagaatctctcaataactaagaaatagcgacttcatttttcagcgzcg
agtgcattattttggctaccatgattcagaagcctcttgccaaaggcccaattttattctgctagttttctc
tgttctttgtacatggcccttgcgctgccctaaccttgaattaacgtggctaaatctcaagaatttaagagc
accgtgactgtgtcctcaggctaggaggaggaaatgggttcacagagtgactggattgtgggtctatgaacttc
ggcagccagcagcaaaagtccaggcatgaataatcaagtgagcagtgaaacatctgtagtggtggagatgttgg
cataactatgaatgatgattcaagagtggttttgatgcatattgaataacatgatgataagtagactcgt

bioRxiv preprint doi: <https://doi.org/10.1101/151111>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tgctaagccttctatgtgaaatacatatttaattctcataataactctagagcagtggttctcgaccggggccg
gttatccccctacccacccccaccctcacccttccaccaggacataacatctggagatatttttgggtgtc
acaatcctgggaatgtatgtgctgatatttagaggttgaggtcagggatgctgctgaacttcgtagaattca
taggagaggtctctcacaacacctatctggcccaaatgtcagtagggctcactatcaagaaaatctgctctag
cagtgctgctcatattatccccatggtgaaatagcaagatgggaagtgcaaagtgggtgcttcggctactctt
ggagcagctttgactttgggtgagaaaagccttttaaaaacaatgtttcttcccactcttcccaccccatgggg
aggtgtgggggttgggtgggttaggcaccaaagcaagatttagaagagttttctgttaggaatttataatggtaa
aggatcaacttcatttccaagctatttatgaggggttatgtttaggaaaagtgctaagcttagagaaggagg
agaaatctgattttattaatgagtgtagccataatggcataatcctggcagaagtcaacttgggttcttagag
ggaggctattatgaaaagaaataccctggaacattccccgggtttggaagggtgagttctaggttcaatgatg
ggaagaatttttagaggtccaagataaaaagggcaagattaaattttgtctctcatgagttctctgggtcagg
tggtgtgaactttgcagacagctctctttaattcactcatacatgctagtctccagctcagcaagggtttg
agagagcaggtgtctgtatgctctggtaagtgaaggcaaaagtgcataaggaggttgggggtccataatggcga
agagaaggagcccttcagtcagagtggtttgaatcttggctctgccatttgccaatcttggaccattggggc
agtgtattaaactcttgaatctcagcttctctctgtaaaatgtgtataacaagagtactaattggattgt
ttgatgattaaatgagttaatgtgtataaagcactcacaacctgggtacatagtaagaccttctcattattat
tatcatcatcaatttttttaacctctttctctgatctgcttacactcaccagcttcagctgctccaaatgg
cttgaagattttttgtttgcccctttgctgtcagttgccatggggaagatccattcattttttcagtcac
caacatatttttgagcatctgctgccctacaggatcctagatatgggggctgcagagatatccaggaacataa
gccttgattaatgggtcagatcagtgctcagcaggggtggcaagtgcaggtttcttttaagtggcatatc
ttaaagggtatgtctctzaacatagctttgtgatggcagcatgatgggtacaaaagcacacacttaagt
tcagtagatctgggttcaaacattgggtgcagtttcttatgggtcgtaactgttcaaacctcagtttcttca
cttctaaaacggtaatgatacaacctacctcacagggttattatgaattaaatactggagatgagatacaca
aaacgtcttgagztacacagtagctgcccaatattgggtgtaagtattataaatctacaagctgtgaattaa
ttttacctctctggatcctgztgatattttctagaccattccacctagtgggggcatttctacctgagtc
ccggtgggtgtcaaatagaatgtcagtggtggtcctcagttgggtagaattgggtgctcatctcaacccgct
actgactatctctgtgatttaccttctccagccttagccttgctacatataaaaatcaagacaataatgtt
tctatctcacagggttgtcctgaggattaaattaaagtaattaatataaaaatgtgcttgtacatattgggc
cctaaataaacagtagctactatttatccttaaagtacaaatggtagtttcagagcttcaaggctgatggct
atttatcttactcatactctttgttttagcttcatttttttcccctaatttcattagwattttcttttctctt
tt
cgatcttgggtcaccocaaacctctgtctcctgggttcaaacagttctcctgcctcagcctcccagtagctg
ggattacaggctcccgcacccatgccagctatttttttgtattttcagtagagatgggggttccacccttt
gaccaggctgggtcttgaactcctgacctcatgatcaaccacctcagcctcccaaagtgcgggattacagg
tgtgagccaccacgcccggcctcataagtattttctaaattttatttacagtcagtcattttaaaaggaaagt
tgtattcctgtctttgttaatatattataagtattttattcagctacaagcttggaaatggcatataattttg
tattctgcttttttcaacttaatatattacatggctaattgatttctgtgtttcataaacattattctgatgatg
catgatataattgttgagtacatgtaccataattgaatcatttccctattgctatgcaattaaagtgtttcca
atattttgcaattataatgtttcaatgaatgaataactttatgcataatagctttttgatatttaagtccag
tttcttaggatgaatttccaggaatagtaattgggcaaatgggataaacatgactcttgaatcagatttgtt
aacattgctttcccaaagggtcactgatttatatttccgtgttcatattcttttaaacagctcatttac
tcaccaaacatttttaagccattatcatgtggttaggttagtaagaagaaagtgaccctaaggggagaagct
tatataaaaatagggtccctgggtgtaccaagtgctgatcacagacacaaagtagctggggaaattgagatgag
ggagtcctgggtcagctgggagaaaagtccattttcatagagtcaggtttttgttctttggcagaaaagaaa
ttgctttcttccccacccccacccccagctttattgaggtataattgacaaaataaaaattgtatatctttaa
gatatgcaatgtgatataatgtatatctcaacttaaaaaataagctacagaataaaaagggtgtttgctatt
aaaaaaaagaaaaggctgaatgtcattcccaagcttggaaatttgagtagttgtgctctttgggattattt
acagaaatattagcaagaccagccccatctttgggtcttgagtactccactgtcagcatgctttcttccagag
agggatccatttgctttatttttcttctgtgtgocgtctatgcaaactattcttgatagttttatggta
acagtggttttttgttccatgagatzaatttatcatgctcatttgtggaaaatttagaaaagacaggaaagt
attaaaazacatcacttttttttttttttttttttttttttttttttaagzagacagagcttctgctctgcgc
caggccggagtgtagtggtgatctcagctcacagcaacctccgcttcccagggttaagtgttctcctgc
ctcagcctcccaagtagctgggagtagcaggtacaggtacaggtacaggtacaggtacaggtacaggtacaggt
atgggggtttcaccatgttggccaggctgggtctcaaactcctgacctcaggtgatccgctgcttggcctcg
caaagttctgggattataggcaggagccactgcgccagccacacctacgttcttatcatcactagatcacca
ctgtcattatcttgtgtatttcttctgcccagctcactctgatcatgagtggtgagtgagtgagtgagtgag
tctcgggtcactgcaacctaggccttctgggttcagagtgattctcctgccttagcctcctgggttcaagtga

ttctcttgcccttggcctcccaagtagctgggattacaggcatacaccccatgccatctaatttttgtatt
 ttttagtagacacagcgcttctactaaaatttttgtatttttttagtagagatgggggttccaccatgttggccaggc
 tgggtctccaactcctgacctcaggtgatccgcctgccttggcctcacaaagtgattacaggcatgagccact
 gcatccatcgccaaaaagattttttaaaagagtttaatgtagaacatatacaaaggtctttggaaataaaaa
 acagtttttttaaaatatcagaaataaaacaacaaataaataaataaaaaacacccaaaacaatctgaa
 gcacgagcacctagcagaaaggttcaattatgatctattcatagagtggaaatcaagtagacattacagga
 catgttttaagattatattttatgtcatgggaaatgctctccagtgatgatgttaaatgaaaaaacagaata
 caaaagtatatatgctgcatagtctcaatattgtagagaaaaaatatttttatgtatgcatgaaaaagac
 aaaagatgttaacagagatccattgttacttccagtttactagggattgtctctgggaggttaggattaaaggtg
 atttataattttaccttttttaactttttctgtatttttttatttttcaaattttccataaaaaataaggacttg
 aagatcaagaaaaaaatttctgcttggctcagtgagtggtcagcctgtaatccagcagtttgggagcc
 ctaggggagaggatcacttgaacccaagagtttgacgttccagtgagctatgatctccggatcgtaccgcct
 ggacgatggagcaagaccctgtctcaaaaaaaaaaattcttggctttttttttttttgtttgttttgagacgga
 gtctctctctgttggccagctggagtagcagtggcacaatctcagctcaccgcaacctctgcctcctgggtt
 caagcgattctcttgcctcagcctcccaagtagctgggattccatgcaccaccactatgccagctacttt
 tttgtattttcagtagagacaggggttccaccatgttggccaggctgggtctcgaattcctgacctcagctgat
 ccaccggccttggcctcccaagtgctgggattacaggcatgagccactgtgccagcccaatcttttgctt
 tttttaaaaaaagaagacaaaaagggattttataaccagattatcttggctgtgtgactctgaagccacagt
 tgtaagttataattactctgaaacacaagggcctgtgactcttttgggctcttgggtgtttatcttgattac
 aacgttggaaatatagaaatgaaaggaatgggagaggtgatagacttcaggcagtgtaactagttgtctgaac
 actactggctcaattatattgtgtctagtgtattccatcttgtccgtctgctaatttatcgctggtaactc
 actgaggcaggggttttcccttggagaaacctcattgttttaaccagtgatcatgcttgtttagaagttcaa
 tgatcttttttaactcatcggaagatgatgaccagacctggacagatggggaaggactttgcactctctct
 ttacagtcctgagtgacacaggtcaatatggaactatgtgtgaattttcattgtctttgagagccctcttc
 tctgccccatagggagcagcttgtgtgtgaattagaggagcaaggggtgtgtgtatttagcacagcaggttg
 gcttggtcctctcctctcaacatagtaccacatacctggcactatgctaaggctgggaatgcagacagatg
 ggtgcctgcttccagagtgctcaatgtgtgaggaagccagcaacagaaacagatgatttcaggagctccag
 gaaaatgctacaggagaggtgtgctgggttactggagtagcacaggaggagggttctagctcaggctgag
 attttagtaaaggaaattatgccacgatgaatcctgaagaatgaatagaagtgaaccagataaaagcacgata
 ggaagcatcttcccttacctaagggaagacacagaggtatatggaatggatgtttaaagggttgggactcca
 aacagttctgtttaaagcttagagagtggtgggagagactggagaagttgatttaattagtaaatgaagttgtc
 tgtggatttccagatcccagtggttggatattccatatttttttaaatttacagtgttctatcttattt
 cccactcag**TGTCAGCTGCTGCTGGAAGTGGCCTGGCCTCTATTTATCTTCCTGATCCTGATCTCTGTTCCGG**
CTGAGCTACCCACCTATGAACAACATGAATgtaagtaactgtggatgttgcctgagactcaccaatggcag
 ggaaaatccaggcaattaacgtgggctaaattggacttttccaaagatgctgtcttgggaaacatcacaca
 tgctttggatcagzaaaacctaggcttctaatttgttgataaggcatgaactcaggagactgttttcagtc
 tagtgaatgggtgataattgtaattataacagtagacaacatctcttttacacattttaaatcatgaaaatag
 aataaccttactgataatttttagaaagtgggtgattaaaagcacatttaagataatgccttaacacctagtct
 tttccatagcatgatgtcttaattcacacattgcaaatcatggaacacagaattttaagcagcatttgtgta
 gaacttctcagtttactaatattatttttatttttctcataacaaccttgaatagaactcagatcatctg
 tcaatcatgtattttgataacagcctttacagtgagcatagaaaatacagtagtggtgcataacaacacaggtc
 cagatgtcaggttatctgggtatgaattctgggtcagcattcactaagcatatgacctggacaagtgtatt
 taagtttcttttaaacagagaatagtaataacctacctatattattattgtcagtgatcatcttacaatca
 cagttcttctcttagggctgggctcagtggttgattgacactgcagaaatggccagatctaaggatcaac
 atttacgtagctgggaaatgtagctgggacttcagtttactgcctagtgttttctaccactaagcag
 ctcagttccataccctacgagacccacaagcttatgagatactgttcttccaggaaagcagtggggcccaggg
 ccaccttttaattgtgtttcttggcctgggtcccatctttctcacaatatatagcaacagttatttacttgc
 gattttctaatgcacatcacacatagtcataattaaacacacacacacacacacacacacacacacccctc
 aagaaacattttctgagacgtgatttctgatttcatcaaaaaagaaaagagcgggcccaggccagtgaggaa
 gtcaaggtgggtggatcacttgaggtcaggagtttgaaccagcctggccaacacggtggaacctcgtctct
 actaaaaatacaaaaattagccaggcgtgggtggcgcacacctgtaatccagctactggggaggtgaggga
 ggagaattgcttcaacctgcgagggtgaggttgtagtgagccgagattgcgccattgcactccagcctgggc
 aacagagtgcagactctgtctcaaaaaaaaaaaaaaaaaaaaaaagcataaactgaaatttatatgcaatttat
 atgctgtgagataattctgttttctcttttggaaaccccaagagatttttttgattgatgagcaaatatcat
 tttagattttatttaagcattatgccaagcaccactgaagtataagtttcaagggcaactcagtttttttca
 tctactagacgaatgattttctggaatgattacaagcaggcaazgatgggtgzztagtggaatagcaaatgt
 cttcggcacagacaagtgggggtttgtttgtatcctgcctctgccttcaccgaggttgtgtatcttgggca

[illegible]

bioRxiv preprint doi: <https://doi.org/10.1101/151111>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

gtgtggggagccatgggagtggttagggccagcctgtggaggacctgggagccaggctgagttctatgcac
ttggcagtcacttctgtaaagcagcagaggcagttggcctagctaaagcctttcgccctttcttgaccctt
tacag**TGTGGCTCGCCTGTTCTCAGATGCTCGGAGGCTTCTTTTATACAGCCAGAAAGACACCAGCATGAAG**
GACATGCGCAAAGTTCTGAGAACATTACAGCAGATCAAGAAATCCAGCTCAAgtaagtaaaaccttctctg
catccgtttataattggaaattgacctgcaccagggaagagagtagccagggtgtctggggcttgttccca
ttagatcttccccaaaggggtttttctccttgggtggctggcctgtggggccctctccaggaggcattgggtga
agaaactaggggagctgggtggccacagacagtgatgtactaatcttctctgggaagacagaagaaaagtcctc
cagggagaataactacagacttggccttagggacagctaggggtgcagattgctgccaaactgcattttttct
gaagtggccatatgggtgcagtgatggatttatagacagagattttctgtgcataaagagcaattacag
ttgtaagttgatattggataagtgaaagtttaagcacttctttctaaaaagagaatgcaattcattttccctta
atcatttcaatttagtctgatzgggcatttgaacttgttgtctttaaaaagtgaaatctttacctctgatctg
gtaagtatccaggcaatttcttgtgtgccaccaggagggtatctggggagtgaggcattttctgactgaggca
ttggctgccatagcatcagagcagccttccaggcagtgggcctggcaaggggacagaggctgggtgggagcagc
tggctgagtgagccagtaatggcatgtgcattggctgttagagaatgtagaagcaataatgaagccgataaa
agctggctctgcattttattattatcatgcgcgggtgggttctaaacaatgtcagtgataaattactcctcccc
atcatggaccaatggctgccactgctccagggaagtgctttttattccgtttgggtgtttagggagggatgga
gttggctggcctttgctgaaaggcctaccagtttgttttctatttggcaaaagaagaatgataaagtttata
gagtttaaaccagactcagatttgagtttttttttttttttttttttttttttttttaaggctacagaactgtg
ctttccttgggcagtaaaagaggcaatgggcaatgtgggacctgatzzgacaaazagggaaazzzgzzzzczzaz
zzazzzazzzgctgtcttaggggtggcatggaggagggtgctgcttcacagcagagagagggtatggctgtgctt
ggagtgtccacttagacaactcctggctgtgcagccaggccatcgagatgctgtttccttgacctgcaggctc
ctggctcttgacatggatgtttcttctgggtgcaggagacagaaaggtagcaacaacccctgatcaaagcctc
agtccttcttatttactggagagccctgctgattgaccagaggcacagctggggatatttcttctacctc
tgtagcaagagacagcgtgggtgcagaggaaagtgctagcatacattaccctgtggctgcagctgatttgtgaa
taggttagttagcacccttcagccacttcttcttacctgttaatgagataaaacatgttaattgtctaaaaa
cagtatttggcacataggaagcacttagtgaaatgaattatgattttttttggagtggtgacatctcaacc
aagccatttaacccctcagccttzaacttctcaactataaaatagcagctaacttgaaatgtaaactataa
aacctaattgtagtatctggcacatagtagattcccaataaatgagagccagttattctttctaaagacagtgat
gcatttctgagcacctggccttgttcttctgccttgaatttatgcagcagttgaaatagactggctgatgg
gggttaagttgtcaagcagacttctgatcttagtgaggagactgccttaaaacaacactaatttcttcttt
cttttcttttcttttcttttaagacagagcctcgctctgtcaccaggctggagtgcagtgggcagctcttgg
ctcactgcagcctctgcctcctgggttcaagcgatttctcctgcctcagcctcctgagtagctgggactatag
gcatgctccaccatgcccactaatttttgtatttttagtagagatgagatttcaccatggtggccaggatg
gtctcgatctcctaacctcgtgatctgcccgcctaggccttccaaagtgtgggattacaggcatgagccac
catgcctggccttctttgagaagctggagacatgagtttaagtggtgaagaagccaaatctgtatctaaaaac
cctacagtagtgagcagagctctgaggagagaaggtcccttagattttgagtgattattatgtcagtgctt
gttttacatctctctgttcacgcagtagtccccctttctgccttgacagctgtttcttaaattctttcttct
tttgcttgtcttgagcacaaaacaggccttcagtaggggggaaatgcacagaaacactgccttttcttaca
ggaaatcagtaacttttactgattttgttttttatttattttttttttttttttttttttttttttttt
tttttttttttagagacaggggtctcttctgttaccaggctagagtgagtgagtggtggccttagagctcactga
gctcactgcagcctcgaattcctgggctcaagtgatctcctgccttagcctcccgaagggtgggtagaca
ggcatgagccactgcacctggccaacttttgtgatttgcaatagcactcttgtcaatttcggagagaagc
tgagactggcatatgtcagtagtgatccccacttagagacctgtgtttatctgcactgacaccccatcacag
catgatgagcttggccctcctgtgctgtctctccagggtgggaggatccttgaagctgatctgggttggga
gctttgtcctcattcacctcctttaccacacacccaccttccaggggcggggatctaccactcactaagtag
cccattctgggtgttgacagctctaattgttagaaaatattcaccacccctgttatgctttctagagaacaag
tctaattctgttttcttgaatagtcgaagacagctctcatgtttttcttccctgttttcccaaagtccta
tgatttttaggcaaaatggcctccttctcttattggaatgttttcttccctcttctgctctcctctg
gttgtgtttcagtagtctgtgtgtgttcttgaagtttactggaaattatgaaagtattctggcacagaggag
gaagggtttttgcctcccttgttctgagtgctacatttccgttaatgcagtgctgagattgtattaggcatt
ttggcattcacgtcaccttgttgactcatattccatgtgcactcaacaaaaattgtgattattttaaatagg
cagaattgcaagttacgtgttctccatttcttgggtgtattgttggctttttgaaactaaagggaaaaatgtc
tttttctgttttacatgttttagattccctatgctatcctatcctcccaaaaccatttttagattctgatttt
gccatgtattatctgatactccttctcgtcatctagagatgtgataaacaactctcttggcctcattc
cagtcacgcataactgtgggacaaaagacttgaagcttggatcagtcagtgaggagactaaccacccctgtaga
cccttttttctcctaactataaaatagcagctaacttgaaatgtaaactataaaacctaattgtagtatctggc
acatagtagattcccaataaatgagagccagttattcttctaaagacagtgatgcatttctgagcacctggcct

tgttcttctgccttgcaatttatgcagcagttgaaatagactggctgatgggggtaagttgtcaagcagact
ttctgatcttagtggaggagactgccttaaacaccactaatttccttttcttttcttttcttttctttt
tttttaagacagagcctcgctctgtcaccaggctggagtgcagtggcgcaatcttggctcactgcagcct
ctgcctcctgggttcaagtgattcttgattctgtagacactaccactcaggcctatattgtaatcagtgcgtg
ggccactgggctcctgcttctgtgatccagttgggaagtttatcttgttcttcccttcagcttgcctctgct
aaattcgctggactatacacaggtgatttgtagatatggggatctctactcaaatactctcatgatttcctt
ggctagagcatcattttatttccacttattggaagagaccttagagaccagttagttcatttatagataaat
tagttgattctgtcattcaacccttatatatattgagcgtctcctatatgccaatcactgttctaaagtgccgag
acacagaggtgtccaaaacaaatatggccctctccatatggaatttctattctagagaagaatctgaccca
gaagggggaagtgactgtcccaagtctacacacacagaggatattctgggaataaatcacggctaaac
ccccctgctgctccaggcagttctcctctacagtgccttattgtgctgttttaataatcttcaactggga
agaactcccatttcaggaattaaagcgtggacaaatctttaaattatccttgaaatcatcctaataagaaat
ccaaggaggaagtcttacaggggtgcctcaccacttttctcatcactggaacttttagacattttattatt
ttcttctaaaccagagtacaggcacacaagttgagtggtgtggtggctaaattaattaatgtttgcaaggc
agtgtgagaagcattcattcatcttaaatacctatggtgactgcaactcagatgtaaaaaattggataaatcc
tcagaaaccctagggaagtgacatgtctgtattttgtctctgtgagatacagactggcagagataagtgtt
tctctgggtgagttttgtgtggtatctgggatgatttttaggcagtagctggatgagaacttttaattttaac
ccacatccaattgcaatttcatggaaattattgcttaggaggatgttcaacaggaaaaatataattaaagtt
aattcaaaagaaacattttctgtgaatatggtaaaacttgtgagagtagtttgtaaattgattgaagattgga
aaacattgggtataagagttagtggtgggttttgtattaagattcattttgggaagaaatccatgctgcaccc
ctcatgaagtgtgaactttgggcatgtgttgattcttctggcccagagtttacctgaagattagctgcctt
gagggctactgagcattaaattagatgatgtctgtggatgactgatagtgaaagctcatagcccacagttgac
acataataaattcgagttgctttgcttcccttctgttccctggctgactgtttggcctttggcacttgttcg
gctctctgggcttaagtttcttggccttgaaattggaactttcttgggtgaaacaaccagaaaaatgcttc
agcccagaaacttggtcagtagcttggatgggggatcacctgggactaccaaggatgtgggctgtctgctag
actatagccccctgaggggaaggtgggcgccttgcctatggttccatcctaagccccagcacagtagtgggt
gcatgggtgagccatttagtgaatcttgtggaatgaaggtgggagaaaaataaaatcacctgtacttcacaggg
attgtgagggctcaagtaaaagtgtttaaaaaattgtattatagtttattcccttgtgttagcccaggctc
aacagagcctacgaataataatgatgacagaagttcttcaaaaagtcttggccttcttcttccaaaaatt
gccccccagagcttctggaagggcagccatgaaccacagaggcctaaagtagatttactgggaagctaaaaa
tatttactttatttttcatagctcctttcaaggtcctctctgggggtcttagcaatatgtttacacagtggt
atgtttttgtaaggtttgcaaaagtaagattttttaaataactatcttgttttaaaagagagccccctac
caacttgtgtcagcctcaggccccccacctgcctctgctcctgccagggcatggtggggcaagaagcactgct
ccccctccaaagcttccctccttggcctggagtcacctcactccccactccaagccacctgccatcgctgt
gccccctctctggtgaatctggcattcttaggtgggctgagaagcagactggcccaagctaaggccttctg
atgggggttgttgcctgctgagaatcatgactgggtgggagaaggaggtgaccttttgctgtcttatttttac
tgtgtatttcccttttcagctacttaaaatgtattgcttagtgatacctaattgggttcattagcctgcttcc
actgaacatttccgctcaggcatccacttgggtcccaaggcctgctcctctcccatattctgaaatctggact
acagactctcattcaactccaggttgcactgtggacacagtcctcctcttgagcaggtacctccttgcagtgg
ttgggagctcctacttggctcatagtgggaagtgcatactgctggagctgaagcctcttgcctccccggat
agggcgtcctcacatccccctctgagaagttccccagcttccctctgttccccgtttccacacttagcgaggc
tcttgtccactgctacatcccccatagccagtcctcctcagccttgccattgcttatgctgggtctggaacaat
tcctagacttgtggggcatctggggaagttctccatcttttttttagctggcatgaccaagtgggtgtgggca
gggctgtggattctatggtgtgggtggaagccaggtagcctctctctactgtacatggaactcagcaacttc
tgagtcaagcaagatcttagctctgcaggtgtcttgcctgtccaaagttatggccacaccagtacctttta
actctagaagcccagtaagtgtttgtgggaccgcaaagatcattttctagacctgctgaatatgcctagaa
cgggtagggatggctttcacgctgttccctagggtgacaagtcacacgtttctgggggtacatacacaccgc
ggctcctgtgaatggcactctccatgagaactgtgatgatttgagttgaatagtgcacagcctacatgggtc
tctgcatggcctggagttccttatcttgccttctccagttaggactagggctgcaactggcctactttggc
tctgacttgggggattctgaaatacctttttttttaagggttgtggagctctctgaagcttataaggattt
tgccaggaaaagataagaaatatcttgggcattttgtcactgtgctggagatgaaccttttgaggacata
tcaccttgttgaggtcaaggggaggaaaggacaggactggcagagagatccggggcagcagcctgccatcc
cgactgagtatggagtttctctctcccttcagctgcacttttgtgtggagtcagtggtcagctgccacttc
ccttatgttcatggcatgaatctggcttgttaggccttcttttttttttttttttttttttttttttttt
ctctgttggccaggatggagtgcatggagtgtagtggtgtgatctcggtcactgcgaccacccgctcctgg
attcaagcgatttctcctgcctcagcctcctgagaagctgggattacaggcgcatgccacaacacctgctaa
tttttaatttttatttatttttttttttcaaggcagaagaatttttcttagtacagaacaaaatggaatc

tctatgtctactctctttctacacagacacagcaacaatctgatttctctatcttttccccacattttcccc
ttttctattcgacaaaactgccattgtcatcatggccattctcaatgagctgctgggtacacctcccagat
ggggcgccggcgccgggtagaggggctcctcacttcccagaagggcgccgggcagaggcgccccccacctcc
cgagcgggcgccgctggccggcgccgggctgccccccacctccctcccgagcggggtggctggccggcgccggg
gctggccccccacctccctcccgagcgggcgccgctggctgccctgctaatzttttgtattttcagtagagatg
gggttttaccatgttggtcaggctggctctcgaactczzczztzgacctgttgatccacctgzgzccctcazgc
zzctcccaaagtgttgagatztacaggcatgagccactgcgcccggctctgttztttttgtttgtttgtttg
ttttgtttttgttttgagacggagtcttactctgttgccaggctggagtgagtgaggagccacgatcttg
gttcactgcaacctctgcctcccgzgggtczaaacaattcttctctgcztzcagctcccgaatcgtggga
ttacazggcacttaccacczagggctggctaattttttgtatttttttagtagzzazgacgggggttttgczcat
gttggccatttgaactcctgacctcttttgaactcctgacctgtttgczcatgttgaactcctgacttca
gzgtgcgttggccztcccaaagtgttgggattacagggtgtgagccacctgcccagcctgtttggcctttct
gatatzgzczctctgactaatcttttttgaaattzagzzztccccagggttatactggattttacttaggg
aaaagggtcatgcztctctggctgtcagzatttactgatagtaactaaggzactcagztgggggtgzzgzac
ctttgattczztzzggtttgatttttgaaaatzcaaaaagacgtgagctccaggggagcagggtggctttgggtg
azcatggcaagatagttggctgtggczagggagttgaggggaagtgggtagaaaattaacatcttgtazaata
ttzzczcctgggaaatataccttctgtgttaagagaacagacttggcagzcczzazzzgzazzzztzzzggcc
taggttcaaactcttggcttgatgcttatcagctgtgttaaccttggataattccatacatctctgtgcctcag
tttctcaaagtgaataacaatagtacctccctcaggactattgtggcaaattaatggacgaataaggggaa
gcacttagtacagtgcctggcccagcataggtaccaggcttgttcttaagctcactgcatttttacaactcat
cataaaatgcaggggatacacacatgaaggagccgaagttcagagagggccaagtaactgcctaaagaagca
cagctggcgaagggcagtaaataggaccgaattccagtttctctgtgctctgttgttattatctctaaagag
agcagctctgagtagccagaagcttccctaaagtccaggacatggggcatgggctggctgggatgagaaag
gagacaagagggcttctgaaagaaatgccagattcactccacttctggcttcaggcacccgatggaatgttt
cccaaggcccatctagaaagaacatcctgtgactcacagccacttctcattttctgtctgaacccccctcacc
cattcaggcagctgctaaagttgaggttacagcctcagactatattttctgtccttgtggaaccccagtggtg
tcacttgttgggagatctgggtgatacatgtgtcaacattatgtcatcaaatggaaattctttgaaactctt
taggtgattgcaattcacgttctgtatgtatgcacttgtcaaaagttttgatttgaggccttagaattttat
atttggaacctttccactaccatgagttttccagacctgtcaaaagccaggctgcactctcagaaaccagggt
ctttgatttccatccaggggcaagggcctgggcccagctgggctgtaagcagggtgggggtggggagcaacgct
gcactgcaatgttgaaatattacttgaactaaatcaaatcaaagatcagctttactcagacaagaatagaaa
acacaattgcattcagattacagaatagtgtgtatccccacaaatatcagactgccttataaaagttttgaat
tgtaaacatcaagaacagtggtgcgtgtctctgtctttccagcataagggttatttattctgtgggtggca
aagagcaatttgggagtcagtttgtttctcattgaaagctttccattttctggctcctcttgtcactggtgc
attgaggcaccaaaggcaatctcagtgcgacactattcaacagactaagttgcaccggataatgataccat
ttacatttttcatatattattacattgaaggcttcaaacagcactagcagggggcaaattgggtattattatc
tccatttctattgatgaggaaaccgaggctcgaaagggtaatgtctgttgttccaagattaaagagtaaagttg
tacgttgaatcggggctgactcctaggcttagcattttctccccacactatgctgccatgttgcttattcc
aacattaggaagcataggtgccatccccagcttttgaggccaatatcacgatgaagcatttttaaaacatct
cattaaattgctgatatagtggaagaaccaaagctttgcagtc aaagctgtttgggttcaaatttaccact
tgtttgttctatgacctgagcaagatattctccagatctgtttcctcatttgaaaaatgggaataataata
cgtttctttataggctgttcttaaagattctggaaaataatgctaatagtgtgcctaatgcttggtaaatat
gagtcactttctgtgcccacaaagcactactatgtcccttaataaatttgttaatttttaaaagttagaaa
aaaattaaactattatacattgtgtatgttaattcttccctagaccagccttaggaagaatctcatcccca
acttgtaaaactcatctttttccggttctttttgtgcctggacttctcagggcctgcaggctgattctagtc
catgttgtgtggtgtttgaagtgtctgggtcccttttttcagttagagaccagctcatccttgggaactgaat
gcctcaaactctctttttctttttctctcttcccttctgtttgatgtagtcttttctgtttctggactctgt
ttcttcatacttctatctcttaccttcttttctactccttttgtcttccagctgtcctctctcatttttctg
cctctctgggtcttcaggtagagttttcatctcagctatcttcttgcctttttctgatgttggttctttgtttc
ttcttctcattctgttcagggtccaaaattcatttgggtcaatgttatgtcttagtggtatcttccatttct
tctgagcttcaaagccaggctgactgtgcccttccacgcctggccagtggtgaccaggacatcctttccctc
tggggctgcactggctcttttgggggaattgttaccattcagggtcttcaacccctcattctagggacttccag
taacttctcccaccttccctcttctcagtaagacatgggtattgtcttattctgtttctgctgctagaagaa
aattctcgagactgagtaattttatacacaatagaaatttacttctcacagttctggaggctgggaaatccaa
aatcaagggtgttggcagggtttgggtgtctgggtgagggctgtctctgtcttaagatgggtgccttgttgctg
atccttaggaggggacaaacgcctatttctcaactggcagaagggaactgaagcctctccctcaagccctttt
ataggggtctcattgtcagqccctctaaqcccaaagccaaqccatcgatccctgtgacttgcacatatagc

bioRxiv preprint doi: <https://doi.org/10.1101/151111>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cccagatggcctgaagtaactgaagaatcacaaaagaagtgaaggccctgcctcgcttaactgatgacg
ttccaccattgtgatttgttcctgccccacctaactgagtgattaaccctgtgaatttcttctcctggct
cagaagctccccactgagcaccttgtgacccctgcccctgcccaccagagaacaaccccccttgactgta
atthttccattaccttcccaaatcctataaaaacggccccacccctatctccctttgctgactctcttttcgga
ctcagcccacctgagccagggtgaaaaaacagctttattgctcacacaaagcctggttgggtggtctcttca
cacagacgcacatgaaactcatgggtatctgtgagggggccctaattgctattgggtgagggcactgctcacataa
cttaatcaccccttaaaggccctgcctcttaatactgtcacattgggtgatgaagtttcaatgtatgaacttt
gaggggggacacgttcaaatcatagcaggtgtgctattgccttactagcaaagtaatctgggggaagatgag
taatgtctcgttccacccctgattccagtgctgactgattctggagtgccagagaggggtggaggctcacgc
ctctgctcacccgacccctctggccatctcctcttagaatgcaagggcagggatttgggttacacagcgcctct
tttgttataggctagctcctgccttcaaggagcactgagaaaaatattatccctttgaaccacaacactagt
atthttgggtactgtagccattcagaagtttgttgaactaatacaagtttatattatttgtaaaatactagaa
ggaatgtgtggttcttaaagttatgggtattcccattgggtgaaagagaagazctctgggttctaatgcctt
ggaaggttcagggagcataagcctaaatatcctggtttctcttgggaattcttcacagcttgtcatctatga
ctttatccacatttatttgaacatgtttaattttcagcttgtattttctcaaggagtagtgtgtgagtcagc
attcagagtacttaacacacttagtgtcattagtgataacttaaccacacgtaggaagctagacctcca
tttacagaaaaatggaggctcagagagcctaactgactttcccaaggccacataactggaaaattctggaac
tgagttttttcaggtatgtcttggtatccctcggtgagccttctctgtagagaaggaagtgctcctgggtt
ggaaagcccaggccgggatgaacagcatgggtgtccttaggctgtgtgtaccacacacccctggcctactct
cccctgctctaaggagtacctgaaggagcagctgcatcacccgcctcaccccttctcctcagcacccaac
aaccaggtcttcccagagagtgctcagctattgggtgaagcaggcatcgtctattcttcaagaagacagcaga
gcatcaggaaaacaaaacaccaaataataccataaagacactggagtgatttgttaccagcccttgatata
acagagaagtatthtcaaggctctatttgttaaggctggacaagctcaggctgattaaaaccacaggttagaat
ttgtctcccgctcccaggccctgctcagttgacagctcctagcaggctgggaggccccacccccacccccc
ccaggggtccattgaagaagaccatctgggctgggagctgtgattttccctacttctcctttttttt
catttgccctcctccacccacttctaaagcagctgattgtctgggttcttttcatgttgagcaaaagagaag
acagatcttgaaacctgagctggggactggagggcactgaccgagagtaaccccgctcgggtcagctcagct
cagcaagggtcttccctcctccgcttctcagctgctctcctcctccttctccttcttctccagcc
tcccccaagctccttccccctctcttccctctcctcctcctccaccttcttctcctctatgctcagag
aactgaaggaaactgaatggcctggaaagcaggacctcgccccaccccgctcataccctccgctgca
cgtgcaaatgtgttctcagcttgggtcccagaggcctggacctgggtcccagaggctccagctgggttccccg
gcttggtcccagaggcctggatgtgtggcaagaagctgaggggtctgcttcttctccagcccacaaaca
cacatgtcacctgttttttgttctttttcctttccttttgccccactttacagaagctcttggttaatttggt
gggtcactgtttgctcctccaggcaaaaccaaggaggctgagctcagcaggctcaagtttttgggtctgtattg
aagagactgtgggtgggagagccctcttgaccttctgctggatagtgactttctcttctgctggcagcagtag
cccagctccgcatcctgtgacagaggctcataccttccagttaccgagcgaacacttggagcaagtcagctt
tttccctgaaacagcacgttccagttcttgatcttgggttgcaacaacactgtgaatgttattagcccc
gttttttttgtttgtttgttttttgagacagagtgctcgtcttgttgccaggctggattgcaatggcgtga
tctcggtcactgcaacctctgcctccaagcaattctcctgcctcagcttcttgagtgtgggttaccagg
catgcgccaccacccggctaattttgtatttttagtagagacgggggttcttcatgttgggtcaggctgggt
ctcgaacttccgacctcagttgatccaccgctcttgccctcccaagtgctgtattagccccattttatatac
caaggaaactgaggcttagggagatttaaataggggggaaagtatagcttcatgattacacagtcaggaaat
ggcagagctgactcttgaattcagacctgtgtatgctgtatccagtgctgattttgttttgcctgtgttcc
tttcagtgctcaggttaagaagacctcagagatcatcttctccaaccacacattttactgattataaaatcca
gggtccgcattaaactcattggccaaagttaagtggctaaatgactctcctgttcagggaagtattccctgtctt
gtgtggaatgattgggtcttctcgtccaggctcatcttccacctatattccatccttgagtgactgggtggaatg
atcattcctgggtgattccctgtaaaatactagattttatttattctttcataaagttatgagtagacaaatca
gtcagcttcaacctttcctttaccttatagatagtatcagcagaaagaactcaataattctgtcttaggaac
tcagccaggctaatatgtattataatacacacatacatatgcacacacatacacacatgtgcataccttaga
tcacaaattcaacatctcaacttagccgccaagccacztctgtgtcttgtctattctgaatatgcttttaa
gggttaaattagccatgaatctgttcagatcagtcattcctgaaaatgatgatgatttagccatcctttacc
gggcaccttactcagtgcccggtattatactaggctctgcagggtacatttgcctcatttaacacctagctcag
ccaatgaggtggatgccattgtcatctcattttatcatttaatacaactcagccttctctcctgttccc
cactctgctctagctgccactcactgactctcctcttgatctgggtgacgtgggtgagtggttttagca
gaactcttgtctcctcttacttgggtttgttcaagttcatgtcctgcttcatcttttccctaactcactgc
aagggtgaattttctctctgcaatcggaattcccaaatcacatcaggtcattaatttattcactcatttg
tattcattcattcatttagttcatccatccatctattcattcattcattccaatttactgagtcctcctctgt

gccaggcactgctggcttttgaattaaaaaccctggccctcaaggaatttctattcttctgggagcaggata
tttcatatgggtgactgczagcttgatctgtgtggtactagtggaaaacataggatgttgggcacaaaagagt
cacaaaacgccgggcgcggtggctcatgcctctaatacctagcactttgggaggccaaggcgagtggatcacc
tgaggtcaggagttcaagaccagcctggccaacatgatgaaacccaactctactaaaaatacaaaaaggaa
taatagctgggctgatggcgggtgcctgtaatcccaacaactcgggaggtgaggcaggagaattgcttga
acctggggctttggaggttgcaatgagccaagagcgggccacttcaactccagcctgggtgaaagggtgaaac
tctgtattaaaaaaaaaaaaagtcacaaaaagggctggccacctaacccagcttgaggatgggaaggccag
ggcaggcttcttgaggaggtgactcatgggcatgtcatgaggatggggtaaaggagaaggtatgtttca
gacatctggctttcttaatatcttggctttccaccatttttctgcccagcaacataggggagaagactgag
accagcagatacaaaagccactgtactctgtcctcaccctcctttttctcccttctactaccagatc
tgaggttttgagaaatctcttctctaattatctgtctgttcatttcttgggtgagaatctactatgtgcata
gcattatgaccacaaaatgtcctggctcgttgtcatcatgagaggtcttacccttctctccagccacttct
agggatttttgactctgtccttttccagaacttggctccagtctgggttgcctcgccatgaagcacttacagat
aaacctcatcttgggccagtgtctccatttactgtctccttttggcttgccttatccttcttgccttctt
gaattgatttgcctctttatctccctttcagccttgaaagtctcctgaaggcagggaactgtgtcccatct
tttctataaatggcattgttgtacatttggttaaagttagtgaaaaatcttttgataatcacaaattatta
ccatatattgagcacctaccatgatgagatctctgtatcttcatatctttgaaatgtggatactattacctt
cacttaatttggttaagaaaatttgtgctcagggaaggtaaagtcaactttctcacagtcacacagctatttcaca
ggagagtgcagtatcaaatttaggttttctgggtcccaaaccctgatgtttccccacatcattgtttctca
gacttggctgcaatcacctgaaaaattttattaaattcttttttttttttttttttgagacagagtctcgct
ctgtcaccaggtggagtgaatgggtgcatctctgctcagtgcaaccactgcctcccaggttcaagtgtat
cctcctgcttcagcctcccaagtagctgggattacaggcacccaccacactcagctaattttgatatgt
ttagtagagactgggtttcaccaggctgatctcaaactcctgacctcaagtgtatctgccacctcaccagc
caggttttattaaattctgatgcctgggtcttaccctagcagttctcatttagttgctccatgggtgtggcctg
ggctttgggatttttgtgaagccccagctgagtcacgggcagtcagtttgagaaccagtatactacac
catgttgccttctgttttccacacaggttggtgctagtgtgtgagtcacagcctacatgggtggatattccac
gtgatcaggctgcaagtcttctgtagtggaggtcttggccctccctgctgcttccccaccagcgaacct
accacaccatgcacgtgcacagccaaggggtgttgactgtttaatcagtcctctgggttcttctcatcc
gggggattcactgataggaggtggagcttgatcacatccagatgttctcatcttcagagcctgttaacata
acagaagtcttataaacatgctgagcacactcactgggtctggagagtgttgcaggatgacctaggccattg
caggggtggcctgggtcgccctcccacacttctgccacctgcttgggaaggaggagctgttttgaagttc
cctggtgcaatatgatgtattgtctggtctgctcaggaggaagctatggccacctagtcaagtggggtt
agctaatacgtgtattctgtttctgttgagggttccacagaggcctttttgaccttctattttatagataag
gcagtggaggccacaactacatgaaatgactcgacaaataaatggatttagaaccaggtttctgactccag
ggtggtgctttttccatgggtgtacagtgattaatgtctaccttttcacaccagtcctcaactgaagacacc
agcttacaccttcccttctgtttctcccaagaacagaaagtgaccccggtatgtcgctttctgttccctgggaa
ggcagttccagtggttagaagtcctgttcaactcctgggtgtggcctggggatggtcctgacatccctgggc
tcttccctggacctggccagctaaaaggaaatctcctatgatggtactcagatacttttgaaccttgtcagc
cctaataccatctcctagtgttttagtattcagctacccttcaactgggcagtaattctgtgccaggcctgatc
tgggcattgggtggtactaagaacgcataatccctatcctataggcatagtctggttagagacaacatgcaagt
aaaacaatgttcttccctaaactgtggttccacacctcctcccccaacattaaaagtgtgaagggtatgcttattc
aaatgtagattttaggtgctctgcaactctagaccactatttcagaatctctggggagctgggccaagaaact
gcattttcgcatgctccctaaatgaagcttaggtgctctgaggtttgacaactgcagtagagagcctaagtc
taacagtgtagagtcacatgtgatgggaagcatcaggtaggttagcagtttgcaggagcactgattctgaggg
acactaactgggcctaagaacagctactggctgtcatgaggaataactaggagctagccatagaggggtagc
agtgaatcattttctctagcgtatgtaaatcttgcctcaattttattctgtctatataactcaatattactgaagt
ttgcctaaagcagaatacacctggatcatacagcattttatgagagactggctgggctgtcaggccctcctgt
tactttatctctgcatgtgacctcttasstccgcggttaactcctgtcctcattaagcctcacactgtag
ccccattttcagatcaaacctgtttctcttctggtaaatgatttcagtttgcaaagtttgccctctagaggt
tgcttagtgctggccatgtgggctcagttcatgtggtcctgatgagctgggttttattctttattacaaagaa
gttaggtgttaggagagtgggttgggaaggagaagaggttagacagccaaatgagatgagtcagggaacta
tactgtttcgaggtcatagggctcctaccaagcatctggtcagaaacctctcattttggagatcaagaaatt
gaggttcagaaagatgacatgaggtacgcagggaagccaccagacacagcctccaactctagaaactaaaat
cttgattcttagtgctgctttttctgttttggctgactggattgaagccttttctaactgtactcagaggg
cctattatttagggagattccgtatgaaatcctttagcaatcaaatcatttaataagggtgatggtttaaat
atatgttaatgtgttttccctaaagcctggcagacctgggttttgagctttgtatcacatgttttatgtttg
gaatgaaatgagaccatgtctgtgaaggcactttgatatgcgtaatgcactctgccagtggtttgtcaaac

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	
0	0	1	4	9	16	25	36	49	64	81	100	121	144	169	196	225	256	289	324	361	400	441	484	529	576	625	676	729	784	841	900	961	1024	1089	1156	1225	1296	1369	1444	1521	1600	1681	1764	1849	1936	2025	2116	2209	2304	2401	2500	2601	2704	2809	2916	3025	3136	3249	3364	3481	3600	3721	3844	3969	4096	4225	4356	4489	4624	4761	4900	5041	5184	5329	5476	5625	5776	5929	6084	6241	6400	6561	6724	6889	7056	7225	7396	7569	7744	7921	8100	8281	8464	8649	8836	9025	9216	9409	9604	9801	10000

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TTTGACAAGTCTGTGCAATGGATCAAAATCAGAAGAGATGATTCAACTTGGTGACCAAGAAGTTTCTGAGCT
TTGTGGCCTACCAAGGGAGAACTGGCTGCAGCAGAGCGAGTACTTCGTTCCAACATGGACATCCTGAAGCC
AATCCTGgtgagttagacttgctcactggagaaacttcaagcactaatgctttcggaatgtgaggcttttccct
tggaacagcatgactttgtttttagaaaaagtagcggtggctgggagtttgtgatataatttagttcagtggt
attctaagtgttcttagtggttctttcagacttttgggccatctcccaaagggtgaatgggaagaataagctg
ggtgtggctgagtttaagccaaaagttttttgtgcttgtttcaatcagagaagacctgctttttcatgtttt
tactattataataactaagcaagagctcatttgaaaacagagttcttcatatttaaaaaaaaaaagtcttgaa
accattgatgggaagatggatatctattttagtttaaaaacccatcataaagatgacattgtgggctgtcac
agttggaaggccctggaattagatgagaccacactatttagcttacttagtaataacattgcaaagaaaaat
tccgacgaagttttttcagcctaggaatcaatagttcagagaagcactctatgagaataccattcattctt
aaccaaaaaatactgggtgagcctgagcagtttgggtcatcagagtgttttatatagttccagaacaaatatgt
ctctaggtgttctgagagctctgggtgaaattcctctcgctacccccaaacatcatcatttaatatccaggatt
ctggtttttctactcaccagatagattctcttaaaaccagggaagattcctggaggaaggatgtatctggaa
agagatgttcttattataataaaatgaaattgtaatactcttggattttgtgcagcagcaattctttatag
agagttgggtctccagagaattaagaatactcagtttctggacctgttccagatcataccctagaatgt
gaccttagaaacacacttcaggattcatacctttgattgacctcaaaaagtttttgtatcggccagggtgtg
gtggctcacgcctgtaatcccaccactttcggtgcccagggcgagatcacgtgaggtcagcagtttgag
accagcctggccaacatggtgaaacctgtctctacaaaaatacaaaaaatagccgggctgtggtgggca
cctgtaattccagctactcgggagggtgagggcaggaggatcgcttgagcctaggaggtggagggtgcagtga
gctgagatctgtctcactctgttgcccagggtggagtgagtgagtgagtgagtgagtgagtgagtgagtgag
tctctcatgcttaagtattctcatgctcagcctccagtagctgggactacaggcacctgbcaccacgcc
cagctattttttgtattttttagtagacatagggtttcactatgttgcccagctgggtctcgaactcctgagc
tcaagtgtctgcccacctcgggctcccaaagtgttgggattaaaggcatgagtcaccgtgctgggtccca
tgttataattttaagtaaggatatttctctacagggatctttgcaaccctaagtaazctggcctaaaaag
ttagagaagctgactgtgagacatttgagcagctgttgggtctttttgtgctgtgaattagagggctgaa
agggtattatgaatgggtacaaaactttgttacaaaacattttcttggactgttttgggctgcttactgca
tgacaaatgctcaccctttcagctggaatgattgaaattttggaaaagatgggtgtttttagaagacattgt
aatttgttccgggtgctgtgcccattcattccatttcaactctgtttactcattaaacacctattgtgtacac
aaccgggtaaaatccctccactcacacaatgctgaattatactcatagtagaatgactgttttagccctcat
catctgataattaacagctcaggtttcaacctgacagtatctctctgggaggattagcagcgtgacagagt
cagggaatgcaccttcagaacctcagctacactgtgtcccatcctgctgtgttgtgggtgtgtgcttgtgg
atgctgtgggtttatgaccaggtattgattaaggtgggtactaccaggtgctttctgcatactctgggtttgt
ggagcactcaggttctgcttctgccccctctgctgttaccagagacctctcttcaaaatggggctcttgagt
tagagtagaatgagtgatcaggattgttttgtgtaagatgatttctgaggaaggctttaggatgaaatgac
ttccaaacattttgaaatgtgactcttacttattgaaatgaagcagggtccttaattggaatgctgggactgat
acttgatttgcattaaagcagcctttttctattgtctgcttgggttgaaatttcaacatttgtgatggtagatg
gatgtgacatgtgatgacattgcacatgggcagtttaactgtgccaagaagtgcagcagtagcagcaaccgga
gatgcaaagcccaacatgatggggagagaaactcttctttcaatatgtgcttctgtacaaaagtggattt
cacgagagacatatatttgaacatttctccttttgtgtgtgcgtgagtggttccctgtttccagccaagggt
attgtgagtttctcctgggctccttcagaatctgggtgctctggaagcagtggttttggcaacatggggaa
agtagggcagtggtgggaggtcagctgggtctgggttgaatattgcatttgaatattttaccagcattgat
gtcggataaattatttagtccctgtaagcctcagttttctcttctctacatacacataatatatttgactc
tttgttgtgattattgggttacacatatgaagagcctgggtgtggggcctggcacacaatagggtgctcaataaa
tagaagttgataatttaattgacatgagtagtagaattatgtccttgaaaacaattgcgtcaagatagaag
ttttcagccaggcacagtggtcacatctgttgtaatccagcatattgtgggggcccaggcggtgaatca
cttgaggccaggagttcaagaccagcctggccaacgtggtgaaatccctctctactaaaaatacacatatt
tgccaggcaggcggtgggtggcgcacacctgtaatccagctactgaagaggctgaggcacaagaatcgcttga
accaggaggtggaggttgagtgagctgagatcactccactgcattccagccagcgtgacagagtgagact
ctgtctcagaaaaagaaaaaaagatagaagttttcttctgtagatcagtggtagaactcataccaagcgaa
gtggctcctgggtgagttttcagtgaaaaactgcattcttctcagatattgtcaagacttttccccaaaga
ttcttattttatgtctcagtcctgaccttgtgtgaaaattaatactggatgtcagaacgctgttgtgttttta
aagttccctggggttaagagcagtttccattaggtgttctctgcttttacttaaaaaatcttactcatgcat
tgagcaatatattttagttcttattatgtgtcaggtattttctaggagctggactcaactcaaaagatctc
cttttgatgagaacaaagggtgggtggatatataatattatctgtgggataaatgcacttagtcatgaggg
agactgttatggagtgcgctcattgtatttgaactgtgagtttaacaacttctaggaggagctcagggcca
cctggcaggggcttcttttcttctgctcagcaagggtgtattttgtgtgagagtgagtgagtgagtgagga
cttttcttaactttctcttgggtccttccctaaagcagcatgtacctttccagagcaggagaggggacact

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tccctgtctcacagaaacctccaatctgttttggactgcaggaaggagccatagtagtgaggaccagcaaatttt
ggcccgagagattgatttgcctccgattgtacttttttttttttatttgctactataacaagtcaccaaaatact
tagcagttttataaccacatacatatttattttcataaggccagaagccaggatacagtagagctcacctagg
tctacttcttagaggctcacagggtgatatacaaggtgtttgtggggctgtgttttttgggggggctctga
agatgaatctgctttggagctcatccaggatatcagatgaattcagtcctcggtgcagttgtaggactgaagtc
catttcccttgctggctgtcagcaggggctggtttttgcctcctaaaggttgccatcatttcttcttattgcttt
tcatgtgactcctctcccacaagtgggttgagtatctctcacactttaaacctctgacgctcctgtcacat
ctgtctccagtcagagaaaagtctttacttttaagggctcaggtgttttagtttggggccattcacataatcc
aggataatctccttatttttgaggttcataactttaattacatctcacagtcgatttgcctatgaaatgtgca
tattcacagactccaggaataggggtgtggacatctttgcaggggatattcagtcctgctgtagcttcttgat
tggaaggaaatagtttatcatataatttgaagtgttccctcagtcgcttttggttttggaccctctgcagc
tctctgcttttttcttgcctatatttggaaagtgactctaaagcataataagcatcacctatttagggttttta
atgtacaaaaccaaagagaggtgactttgggaggagaacatctctgaactaggtagagacattcatcgaaaa
aaatccatcaagtgtttattgtatgtctgctttataaccagcactgttctaggcactgaagttcgaccaccaa
caaggcagttgtgatgcacttggagctttcattctgatgggcattgagatccagggtgaaggctgagtcctgg
gaatttgaggaattcttttgtaggtcctgggtctacagagtgagagctgtcctcgttccagtttctactgatga
cctctcgaccagctccctcacagcagctctttgccaacacagacactgtgggctgtagtgagggaatgtggg
gttgaaatgagctagggtttgggctctgtccttgactcaccattgcctcagtgatgtgaaaatgggtgtctgatc
cttaaggttggaagtccaggcatgcagatttatctccatctcaataacgtggggaaaaaaaagaagtggttt
ttgtggaaccagtggttccactgtccacgtgtcttccagtggtgcttagcacattcccaccaagtggtatcttg
gttcatgagggaaagaagggaagtgaagggtgccccgaggtcccagaagacaagtaattgtcacagctgagc
tgtgtacagatcgaagaagcagatggataaggagtgaaacaaagtcaccttcttcttgaggggtctttatttta
gcttcttcttgactcttagacaaggaccagaatacagatggggcttgttgttaccttcagcctcatggctt
tttagggctagatacttcagcttgtttacatgcagtcctttaaagcgtctgtgtgggtttttgcaggagagaac
acttgctcgtgttctcctgtctggaacccggacattgttggaagatcagattttgttggctttgtgtga
tttgactgcccacctgtttacttgccttctccagcaagcagcgaatccccatgggttggaatgggtgg
aatgacacactgtgtagatttactcttcagactctatgttcacctcattcttatgggaaaaaagagcacta
gctggtagatatagcagtggtttaatatgactgtcactccatttacacagacattttcaccttagttata
ctatttcttcattaaatattgttgccagatctaagatacaggtttaatttttctctgaattatgtggtag
tagatgtattttaactatgtttagaaacagcaaaaaatgaagcgtttgaatgcgttaaacacatctaatttga
aagttaatatatttaggttggtgatttatttttttaaagattagaatattccttagaaatgtagctttataatt
tcgtatttcaataaaaaatattaaaatgtttccagaggaaatctttactgtcatagaaatcaccagaaaga
gatagcaattactcctgggtgggtgatagtcctttgatttgtgggttacttgttttcagtttgaactaaaacat
actaatcccagcctgagttggattttgcatatgggtcaaagtgaaagtagagatttttgtctactttatgaaga
tatttaaaggacatttgaaatgtttcaatgaacacatttgaacatgcattcccaggaggaaaacaaaattgt
gtattgtgttgaaaatactattccaatatatgtatacccacgtctcattttgtcatagaattcctgaaaaat
ttagatgtagatgagatttttataagttgaaaatatttttcagttgcatttttaattgcacagatgtgttttct
acttctatttggctgtgactcctaattgcattgttaattatcttccagtggttcatttctgtgtagtgataatg
gatttaacaaatgcattcattctttagtaaacatttactgagccctaatacagtgccatgctctggcccaggga
aggtagctatgtcaaggctcagccctctctgtttcattcctgggtggggaatgggtgggacagaccaggaa
aggagcagctgocgtatagagcaggagacgctctgctgggtgtgatacagtcctgatggggcctcctactg
cagtccttccagtcaggaaagaaacatcctctaagtggggacctaagaatggatgtcagccaggtaagaa
caggagtagaatgttccaggttagaggacatagtatatgcattgggtccagagaggagagggaagagaagggtg
gcatttaggggaactgtacctgattgagatagttcagcatggagcagatatagctgtggggagtttgatggg
agaaggaaacgatgagacgtgatgcctgcagaggtgggcaaaggtcggtgaggcaggtcttcttccagagc
catgtgaggatatagggccttatcctcagagtaaaaggagccattaaaggctccacaggctgagtgacttga
tgaggtgaatacgaacacagacatttctgtgaatgtgaatggcctttggagcagaactaagttcatgaggat
ggaaagaatataatgaagccatctcttagaccagaaagaatggggcacaacagcctttatcttctggggc
caatgtcacagaatgccatgcttttaggaaacatgggtatgttgtgattaaacacattttgcagaagtgggtgg
gagcttttgaataataacagtaagcattttgtgcattcttctctgtaatgacattacagttatgatctgaaaat
attgagtcatacatgaattcctgttatcttaactcagaaaaatataagtcctcactaaaggttttattttcct
tcttttcccatcttcttacttctgtataagaaagtcacttgtctcctgggtgcaatggagacctatgtgag
ttcatagccaagagaatgttttgggttagaaaaataatagtaggaattccaagctgtgaattttttactgaa
gctcttgggaataggatttggcaagttttgtctgccttctgtcaagtaagcatgagcaggagagcacagttta
atagcaggtgcagacacatgattctcagacggtattttgtgttctagtttcaaggcatgaattcttctctgg
ggttaatttttattcaaggaagttatctgtctgttagatgtgatagtgctcaggccaacatagattctttac
ccttcttcttctcctgctcacctgtccttctcttttattcttttcattgaattaaaaaagaaaattatgaat

bioRxiv preprint doi: <https://doi.org/10.1101/151111>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

agtttcaacatgaaaaaagggtacagagaataacataaagaacactcctggctgggtgtgggtggctcacgcct
gcaatcccagcactttgggagtgctgaggcagccagatcactggaggtcaggagttggagaccagcttggcca
acatgggtgaaacactgcctctactgaaaatacaagaattagccaggcatgggtggcgtgcacttgtaatccca
gctacgtgagagactgaggcaggagaattgcttgaaccaggagggggagggttgcagtgcagctgagatcaca
ccactgcactctagcctgggtgacagagtgagactccgtcttcaacaaacaaacaaacaaaaagaacactc
ctgtaccatcatccatcattttgccgtgctgactccagggttctatttaagaaataaaacattacagggtacag
ctgatgccacctctgtttccctagctcattcttcagagataactcttgtcttgcagttggatgttttaatacc
tctatatcaztgtatacttacattctatgtataacaatatttgggtactggcctaaatgtgttcacattgtat
aagtgtgcatattggcctgccacttcatttgggaattatgttcttgagatttatcaatgttgatacatgtgga
atctgggttaatttttggccatagttatctatttttataactaaacttttaaaaatccatgcttctagtctttgg
cttattttttcagggttatgggtatgttttggatgcacagaaaagtaaaattaagtcatgagcaaaatatctgg
ataatccaagctttaaacttgatgtagaatttgaatcatgtgtgttttgttaacctgtgatgtcaatccat
gcctgattgtgtaactccaaccaatattcctttgaaaatggaaatttgttttatattgactacagattgcca
tattattagtaaatgctgagcacttaatctcgaataaagaactagtttaaaaatgatttcaacaatggcatt
gactgttctaccttattactcatgggtgggttcagccaatgtttctgttggagacaaaaacaaaaacagtc
aaatataacaagcagtcacacccaacatacagactactgataagaagggtcatatcataagatatggcattga
atttgtgtctgctaattgtaaaaatctgatgccacagcaaaacttaataaggacctatgtttacatccatgct
caattacattcctgggttaaacagtcagcttttaggcctgctgtgtgcctggagtttgcagagtggtggg
gcttttaagagaaggagaataagcttgcctcagagtttaagaaatttaaaactaaaagtcctaaagatgttggg
aaaactattgccccttgaagatgtaaattcatttaagttggagaagaccttttatacaaaacacagacccattc
actgatttgtacccttcaggagacagatgaccggtaattgggtgacaatgggtgaatgttgggtttgggggttt
tagaaacatctgcacttgggtgactactgtatctaattgggtgtgacaaacctggcaccctcatgtgttggcac
catcttgggtcctactcagggccagggtgaaccgagtggtcctctcactgcttcagagctcgaagcagattgt
agtatgcggaaccagacacagaatataccaccaagcacgttggcgcaaaagcataactgggaaggagggtttgt
gaacatgggtgctgggttctcaaacctcagtgctcaaaagagtgctcctgaggattcctggatcacactcttaaa
cttctcattcagtaggtgttagctgggtgagaatctgcattttttttttttttctgagaccgagtgctcact
ctgatgccagggctggagtgatttggcgccatcttgggtcactgcaacctctgcctcccagggtcaagcaat
tttctgcctcagcctctctagtagctgggttacaagcacatgccaccatgctggctaatttttgtactt
ttagtagagaggggggtttcgccgtgttggccagggtcgtcttgaactcctgacctcagggtgatccaccact
ttggcttcccaaagtgttgagattacaggtgtgagccaccatgccagcctgaatctgcataatttaacaagc
accacaggtgattctgatacagtagctcccaaacctcacagtggttagtgaatccagtcatttacaattct
gccatgatttttggatcatattcaagtgagctggtagcatttttagttaatatatttttaaatgaagtcact
tcttttggataattaaatttaattacaaggaagctataccactgctgttaaaaacatcacctgctttaaaga
gaagggtacataatgaatatacatattaagataaagatgtatatgtgtgtgtgtgtgcacataatgaatataca
tacctaccatagggattgagttttccttcagggttttcaactgaaatgtcaactttgaggccagttaatatg
tgtaagatatatgtgtgtatgtatgtctatacatatagacataacgtaaaaacatacatggatgcataatag
tatactatacacacaacctattatgcataatcatgtatatttcatccacttagtattatcttztattttgocg
tttggcaaatgctcagtaaaagaaaagggttagaaggggagaaaaggcattttatcccaagccttcagggaatc
aggatgaggatgtcttcaccttgtgggtggggagtazattatacaattagagacagcacattggagztgtggc
tgatatgctgtgtgatgatagctctagctctctgcctagcagaggaaggacatttcaatagaagaaaaagtt
taagaccttgccgagaaaacagagaaaaggatgtttgtctttttaagaagttgaaaaccctgtttgcagacaaa
agccctccagtttttggcagtaaaactttcatgcaagggaagaaaaaggcaggggatgacattgttgacaattg
tgaggaattaccatgtgccaggcactgtgcgaggggctttgtacatatcctctagtttttagtgcttataaaa
actctgtgatatgtgcacagcattttaaactttgctgcatagtcgagaaaatggaaggatggggaatttgag
tcatttggccagggttctatagctaccccagggttcccatgactggagaattgggggcacagggtggcggggga
gagtgagtgcagaagaatcctaacaatcttatttccattgagtccttataaaaagaagtggaattaactaccag
tttttaagtttttcttaaattttaggttatgttgatctggcgcttcttgttttgcctgggtttgttttgttt
ttgctatgctgtcttgaacatctgtcatctttagtcaggcctaacggtaaacacaaaaaacacttacctctata
gctttcaattaaagatctctcagtttgtgtttgtaatagttttccaggcaagtttctcctagggttcggcttct
agtgtgttaaccttttagttataaagtgaacccaagagagaaaagtagaaaacaaaacacctcacctgttttgc
ctcatgaattactctctatggaaggaaacaatcatgaacacctctgcgtatcacagaggcctatctgagtcgt
acgtttaaggagagaccgcttaggtccctttgaggactgtgaatgtgggagtcctgggactctgggtgaagaac
ccgttccagaagagatgaatgagctggacaagttctttcatagaaccttttaggcagggttttcttagaaatgc
acattgaggattatgcttggatattgtgatgatcagaatgatactcaatcccttctgcatttgggaattctct
ttgaaagaaaacatcccaggcagctatttctcagagatagtgagtcacagccacttctagacattttcttgt
gtagtcacattataatttcacagcagtcctctgatatgacaaatgtcaaaatagcccaaccttctctaaact
tcagagatgtctgatatgatattgaataaaacaatgctcatagaacatcaagaaagggtggattttccctgg

[illegible]

aagaagaaaaaagtgtggaaaaagctgtccacctcacctaatttgtgttcttgtgagtggtgctagatgccccc
 tctggagaaaaaaaatccttgtggcctctgacccacctctggagagcctagtcccttctggaggcagaagg
 caaagcttaggacctagagagtgtgtggaccacgccactcacaggaaccagcaggctgtgaggttgaaagcta
 ggcatatggagctttccaggctgggtgcagggcctcgtggcccttccccccccctctgtgctctatagctca
 gtcttcccaggcggtgtgaacacgcagtgcattttccaggaatacagggattttattaatgattttcttgtgaa
 atgttttgaaatacaaagtactctataaatatttcataatagcattgggggtgagaactccacaaagtgccg
 gaatacatttgcattgaagacagaacgctgcctgggtcattgatgcctgttgagtggcagtcacagacactg
 cctagggtttctgactcacgctgttgggactgttctatgcagggcacccctctgtgtggcattaggatttgtg
 cctcaccacacactgttgtagcttttctgtctgtatgatgtagagggcagtgctcaggccatgggtataag
 catctactgccccagggtttaccaaaccgaagcaagttgtgtctcagcgagctccgtgaagcatggagaa
 gttgagtactcagagacatgacgtgacttttcaaaggctgtgaagctgcagagggacatagctagggttcaga
 cttgagtttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttctttt
 gattgcagtggtgcttggctcactgcaacctctgcctcccggttcaagcaatttctcctgcctcagcctccc
 cagtagctgggattacaggcacctgccaccatgcctggccaacatttttgtattttttagtagagatgggg
 tttcaccatgttggccaggctgggtcttgaactcctgacctcaggtgatccaccgcctcgacctcccaaagt
 actgggattacagggtgtgagccactgcaccggccagactcgagttttcatcttaatgctttttcattgc
 ctgacactttactgagaccaagataggggaacttcacatacagtaccttttctcccaaggcggaagagggctg
 ttcaattttctacactagagttcggggagttttagaaatgagtcagttatcgaggatgagagcagttcctgat
 aggtcaaccacaatgagatgtagctgttcagagaaagcattcttttatctataaaactggaagataatcccg
 gtgaaacgaagcccagccccaggggttccactaactccaggctgtgcttctcaaactttagttagcatagga
 atcacctgggcatcttgtgaagctgtagatttgaattctgcaggtcggcagaggggtctcagaatccgcatt
 tccaacaatgtctccagtaatgctgatctgctgcctggaccacagatgggtagccaggttctggcaa
 gctcatcccaaggctttgagatgcagatgcagaaaaatattgttctgggacatggcttttgagaggtcaagaaa
 ataagatgtttctttcttcttctatccccaaccttgcactgcccttttctccctccccctaccctccttct
 tgtccccatccctgacgccag**CTGTTTCAGCATGAGAAGCTGGAGTGACATGCGACAGGAGGTGATGTTTCTG**
ACCAATGTGAACAGCTCCAGCTCCTCCACCCAAATCTACCAGGCTGTGTCTCGTATTGTCTGCGGGCATCC
GAGGGAGGGGGCTGAAGATCAAGTCTCTCAACTGGTATGAGGACAACAACACTACAAAGCCCTCTTTGGAGGC
AATGGCAC'TGAGGAAGATGCTGAAACCTTCTATGACAACCTCTACAAgtgagtggtccatgcagaccccagccc
 gtcccccaaccccatccctcccttagttctggccttggcctgtgtcatctcctccctctgtagcagcgttag
 atgtctacatgccatttggccaccagactgagctcttccctagaggagagagaggttctcttgaatagctacc
 tgtcccagttctctgaatgcagcctggcacatctcaggtgcacagtagtggttatcaatggaatgaatgat
 tgacagccaaccttctggttttctgggggatgtggaaggggtggctccaggggtgacagaatagagataag
 gcagaaggacaaactctgcaagatctcacttatataatggaatatgtgaaggtagaagtgtagttcaca
 tgtagtaaaagtctctgggattcttgatgtacatcgtgatgactatagttagtaaacactgtatagtatacttg
 aaatttgctaagagagtagatccgaagtgttcacactacacaaaaaaggcaactatgaggtgatggatttat
 taacagcttgattgtggtgatccttttacaaaagtatacatatattaaaaacatcacattgtataccttaaata
 tatacaattttttatttgtcagttgtaactcaaaaaagctagaaaagcatttttaaaaaggatgatgtactgg
 tcttaataattaccattgagataagctttataataacataaaaaagaaataacagtaatgataatagcaacaac
 aacaacaacaaagaactaacatttaagtagaatttcttgtgcactgtgcattctgtttaagttatctcattt
 taccctcatgataaacctgcaggggaagattctttaacccacatttcataggtcagagaggttaagtgcct
 tggttagagccacatcagagttaatccacaagagccaggattcaagcccaaatctgcctggatctgtgctct
 ctaagataactgttagtggtggcgtgtgtgttctcacactcagacatttgatctgcctttgtttcccatct
 ttagctgcaaggcagtggttaaagaacctgtgtctccatatccactccccacacttaagcaacttttgtgggc
 ccgtgtgcgtagtgcctcgtggcagcaggatccaattgtcacagttttagggcagtggaactcctttccttga
 aaacttgatgcaggggaacctttctcattttccaaccacaggtgtgtcztttcagacactgagtgzagga
 ggttttgtactttatttgaacacaagaaccttttcttctctggagtaaaagcactccagacattcgcaagttg
 ctttacaaagccttaaaaggatggattgttaggcaactttaattaaatcccatctcctcctctccccagctt
 gcaagttgaccaaggaagccttcatttccatgacagacttaattgtgagggcatcctcattaaaaaaaaa
 aaattctattatctttccagcatatagaagatacttggatatcaaaaatccctgaaaaacttagaatgaatt
 ttaaaaaatcagggatcctgctggataaccaaacccatttgtctgtttacaacttttgtatttgggtttttgt
 taagtgtacataactagtttgtgttaattaaagagaatttttttttttttttttttttgagagggagctcgcctc
 cgttgccaggctggagtgagtgggcgccatctcggtcactgcaagctccgcctcccggttcacgccatt
 ctctgcctcagcctcctgagtagctggaactgcaggtgcccgccaccatgccagctaatttttttttttt
 ttgtattttgagtagagacgggtttcactgtgttagccaggatggtctcgatctcctgacctcgtgatccg
 cctgtctcggtcccaaaagtctgggattacaggcgtgaaccactgcgcctgttgagaatttttttttttt
 ttttttgggagacagagtttctgtcttgttgcctgggtcagagtgagtgacacaaactcgggtcactgcaa
 cctctgacctctgggttcaagcgaattctcctgcctcagcctcatgcctcaccacgcccagctaattttgtat

gatttttttataaagacagcaggtttttcagaccctttggagactccaattcggtagaacaccagagcttcatctctct
ctgtcgaaagctgtgcaggaggttgcaaatgcctctcctttttgctgagtttgacgctgctgtttttccggca
gcacatctgtgcaggcctctgcctcgccctctggzatctgctgattgagcagcggattgatctgtccttc
tctttcgtgttgacccatgtgaggaaccaactggcaagggaazacaagaaatggaaataggcctcctttgcat
catgacctgtacatcctgcaattggaaaagattgtacttttagttgggttaaccagcagcattatttttctaa
actaagcagtaagaaggaattagggttttatgtgggatcaacagactgggtctcaaaagaggaaggtgataga
acacagtggggagggggaggtgcactagaaacagagggcctatgctttcattctggccttgctacttaatag
ctgtgtgaccaaatcttagagacttaacctctctgaacttccattttctcatgtataaaatgggaaatatta
aaggatactcactgggctggtggcttgtgcctgtaatcccagcacttggggaggttgaggtgggaggatcac
ttgagcccaggtgttcaagaccagcccaggcaacatggcaagactctgtctctatgaaaaatataaaat
gccaggtgtggtggtgtgcacctgtagtcttagctacttggtaggctgagatgggaggatcacttgggcttg
ggaggtcaaggctgcggtgagctgtgattccatcactgcactccagcccgggcggcagagcagagacactgaa
tccaaacgataaacaacaaaaggcaaaaaataaaaagtgcctcttttttaggttggtgaggtgaagcaagca
tatacactattcaacatagtaactatataaaggaagtattggtgtttactgttagttaataaccattaaagt
agatgtttcgtatagtggaaagcacatggactctgaattcagactgggtctgactttgagtctcagctccaca
tctagtaataactatgaccaagccctgggttaaaatcatgttttttttttcttcagcctcagctctctcacatat
aaaaatagggacactgtcattttacctcagttttctgtgaggataaaaacaacgacagtgatatgcaagtattt
tgtaaaattttgtagtgctcctcaagatttagttgggtgtttactacttgtaactttctcactggaatggcagAT
GCTGTTGGACAGCAGGGACAATGACCACTTTTGGGAACAGCAGTTGGATGGCTTAGATTGGACAGCCCAAGA
CATCGTGGCGTTTTTGGCCAAGCACCCAGAGGATGTCCAGTCCAGTAATGGTTCTGTGTACACCTGGAGAGA
AGCTTTCAACGAGACTAACCAAGGCAATCCGGACCATATCTCGCTTCATGGAGgtgaatctggttgctgggatc
atttagaaaagacttaacggcttctttctctgagacgttacaataaggttcaggcaggagggaagtttagaa
ataatgtatagtctcatttacaaaactatccctcaagcctaacacaggatttgataacaaaaggcacttaat
aatgttagttgagtgggtgaatgagtaataaactctagcttttagtaaatataactctagcttattctatat
aggctcaagagaatatttctacccattttctcttaggttttctatctcagtgactaatggtagcaaagcat
tcccttaaaaaaggcattatttgtgaaacttatctaaaatcgaattcgggtccaattaaatttttgaaattt
atattaaaaattatatttagtagggatgggtgaagaggtgttttggctcggttggttggttagtgcatagct
cagaattgtctaagaaaaacagaaaagtaagataagatcattgttttaaccttttctctccacaaaatcaata
ataaacatattccctaaattactcttagaatttctctaaattgcagtgaaaaaaccaaactccttcattcttg
gttgaagggttggaaaaactacgttagagaggattagagagagaggatgagcaatcgtgtagtgcagccctgcc
tctagtgtaggatttgtctcagccactgcttggtgtgctctggctgccaacgttctcatgaaggctgttcttc
tatcagTGTGTCAACCTGAACAAGCTAGAACCATAGCAACAGAAGTCTGGCTCATCAACAAGTCCATGGAG
CTGCTGGATGAGAGGAAGTTCTGGGCTGGTATTGTGTTCACTGGAATTACTCCAGGCAGCATTGAGCTGCCC
CATCATGTCAAGTACAAGATCCGAATGGACATTGACAATGTGGAGAGGACAAATAAAATCAAGGATGGgtaa
gtggaatcccacacaccagcctgggtcttggggaggtccagagcacctattatattaggacaagaggtactt
tattttaactaaaaatttggtagaaatttcaacaacaacaaaaaaactcaacttggtgtcatgattttgggtg
aaattggtacatgacttgctggaagggtttttcataggtcataaaaataacagtatcttttgatttagcatttc
tactcaagggaattaattccaggaattttgggtggcaggcacctgtaatcccagctactcgggagggctgagggc
aggagaattgcttgaaccaggaggcagaggttgcagtgagctaagatcgcatattgcactccgcctggg
caataagagtgaactccatctcaaaaaaaagatacaaaaaatgaaaaaggggcttggtgaagggtag
taggggttttgggcaatt
tggttttggttttagGTACTGGGACCCTGGTCTCGAGCTGACCCCTTTGAGGACATGCGGTACGCTCTGGGGG
GGCTTCGCCCTACTTGCAGGATGTGGGTGGAGCAGGCAATCATCAGGGTGCTGACGGGCACCGAGAAGAAAAT
GGTGTCTATATGCAACAGATGCCCTATCCCTGTTACGTTGATGACATgtaagttacctgcaagccactgttt
ttaaccagtttatactgtgccagatgggggtgtatatatgtgtgtgcatgtgcatgcatgtgtgaatgatct
ggaaataagatgccagatgtaagttgtcaacagttgcagccacatgacagacatagatatatgtgcacacac
tagtaaacctctttccttctcatccatggttgccacttttatctttttattttttttttttttttttttttt
gagatg
gagctctgctctgacgcccaggctggagtgcagtggctcgatctcggtcactgcaacctttgcctcccggg
ttcaagctattctcctgcctcagcctccacagtagctgggactacaggctcatgctgccacgcccgggtgac
tttttgatttttagtagagacgaggtttcccatgtttaccagggtagacttcaactcctgagctcaggcaa
tccacctccttggcctcccaaagtgtctgggattacaggtgtgagccactgcaccagcccaccactttaat
tttttacactctacctttttgggtcaaaatttgctcaatctgcaagcttaaaatgtgtcatgacaaacacatg
caagcacatactcacacatagatgcagaaacagcgtctaaacttataaaagcacagtttatgtaaatgtgtg
cacttcttctccctaggtggttaaaccacatttcaaaacaacccaaataaaaactgaacaaagcttcttctct
tagacttttttagaaaacttctcagtgctgagtcagctcaagtctcattgtgggaactatgccttt
ggatgtaatgattttctcttaagacaatgggcggaggtgtagttattgcagacatctgaaattgtaatgttt
cttccagattctggaaattctcttattctctgtggttgggtgggtgggtgggtgggtgggtgggtgggtgggt
gggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggtgggt

[illegible]

TTTGGCTGTGAGTACTTTGCCCTTTTGGAGGAGCAGGGCATTGGAGTGCAGTGGGACAACCTGTTTGAGAGT
CCTGTGGAGGAAGATGGCTTCAATCTCACCACCTTCGGTCTCCATGATGCTGTTTGACACCTTCCTCTATGGG
GTGATGACCTGGTACATTGAGGCTGTCTTTCCAGgtacactgctttgggcatctggttggaatgatgactt
ctagctgatgtcctttcttctgtgctagaatctctgcagtgcagggcttccctgggaagtgggttgggctat
agatctatagtaaacagatagtgccaaggacagggcagctgatgctgaaagtacaattgtcactacttgtacag
cacttggttcttgaaaactgtgtgccaggcagcatzgcataatgtttatacacattgcttcatttaattct
cacaaggcztactzctgaagtagttactataataaccagcaattttcaaatgagagaactgtgactcaaaga
cgtaagtaaccagctttgggtcacacaactgttaaatgttgggtacgtggaggtgaatccacttcggttacac
tgggtcaataagcccaggcgaatcctcccaatgctcaccacattctgtatttctgtgtcctcagagggggta
caactaggagaggttctgttctcctgagtagcaggttgttaataataataactagctctaaggcctgcctg
tgatttaattagcattcaataaaaattcatgttgaaatttctttagtacttcttcttaataataatacatc
tcttgaccaagtgccaagaggaacctgcgttggacagtttcatatgagatcaaattctgagagagcaagat
ttaacctttttgggttcaccttctgatcctcccctaaggaggtatacatgaaatatttattactcctgcctg
aacttctttcattgaatatgcaattttgcagcatgcagattctggatttaaatctgagtccttaacttactg
gctgagggaccttggataggctccttatccctcagtttctcctcatctctaaaatggggatggcacctgccccg
tgggttgttggaggacttacagaggtgcagaatgtacgttgtacatagcaggtttcagcaaatgttagctc
cctctttccccacatccattcaaatctgttcttctccaaaggatgtgtcaaggaggaaatggacctggctg
ggaaacctcagaatactgggatgatgctgagcttgggtcctacactgtgctttgctttcagGCCAGTACGGA
ATTCCCAGGCCCTGGTATTTTCTTGCACCAAGTCTACTGGTTTGGCGAGGAAAGTGATGAGAAGAGCCAC
CCTGGTTCCAACCAGAAGAGAATATCAGAAAgttaagtgtgttgacctcctgctctttctttaacctagtgc
tgctgcctctgctaactgttgggggcaagcgatgtctcctgcctttctaaaagactgtgaaacctccagg
ggcagagaaatcacatgcagtgtccctttccaaatcctcccatgccatttatgtccaatgctgttgacctat
tgggagttcacggtctcgatccctgaggacattttcttctgtgtcttggcttctagaagagtatcttttac
ttgccccctcccaaacacacatttcatgggtctcctaacaagctagaagaaagaggttaagacaagcgtgatt
gtggaaccatagcctcgctgcctgcctgtgacatgggtgacctgtgtatcagcctgtgtgggtgagaccaag
tggctaccacagagctcagcctatgcttcataatgtaatcattaccagatccctaactcctctcttggctct
taactgcagacagagatgtccacagctcatcaaaggctctgccttctgggttcttctgtgcttagagtggctt
cctaaatatttaatagggtcccttttctgccagctctctctgtgcccacccctgattgccttggtaaaagt
atgatgccccttagtgtagcacgcttgctgctgttccctaatactctctcctacctcctctttacacctag
ctcctgtttcagtcacctagaaatgctcacagtcgctggaatatgtcatgttcttccacacctccatgcctt
tgtaggtagctgttctctcacaggagaacttctctcctaacttgcttatcttctcaactcctccttctct
ccaagatctagtctcggatccctccctgagcatccctccttgggttctcaggtagtcagtcactctctgcc
ctgaacttccatggcacgtgaaagaaaatctttttatttttaaaacaattacagactcacaagaagtaataca
aattacatgaggggggtcccttaaacctttcatccagtttcccaatggttagcagcatgtgtaactgtagaa
tagtatcaaaacctgaaattgacataggtacaattcacaaaccttcttcagatttccactagctttatgtgc
gctcatttgtgtgtgtgtgtgcgtatttagttctatgcaattttatcatgtgtgaattcatgtaattactag
ctcagtcagctgcagaaatatctcattgtcacaaagctccttcatgtacccttaaztggccaczagccz
acctcccttcttccctcagttcctgacacctgtcaaccactaatgcgttccctcggtttttacagttttattatt
tctagaatgttacataaatggaaccatacagtaggtatccttttgatactggcttttttttttttttctc
agcagtagtcccttagatctatccaagttgtgtgtgtgtcaacagttcattcctcttctcagctgtagtagtgt
ccctgggaggggtgtatcacagttccatggcatttttagatgtatttttaaacagcttccagcatcctcta
ttttaattgttcatcaagtcctttttcccaatagactctgaatgctcctttatcatcgtattcccatcacca
acatcagtagcccaaataggccctaaataaacatttatagcctcctgcctgcctgagaaaccagggtggacat
ggagagaaggcacttctgaaagttcaagcgcagtgcsctgtgtccttacactccactcctcagtgctttctg
tgggttcatttctgtcttctcctgtcacagTCTGCATGGAGGAGGAACCCACCCACTTGAAGCTGGGCGT
GTCCATTGAGAACCTGGTAAAAGTCTACCGAGATGGGATGAAGGTGGCTGTGATGGCCTGGCACTGAATTT
TTATGAGGGCCAGATCACCTCCTTCTGGGCCACAATGGAGCGGGGAAGACGACCACCATgtaagaagaggg
tgtggttcccgagaatcagccacagggaggttctgcagtagagttagaattttataccttaggaaacctatg
ctgatccctgggccaagggaaggagcacatgaggagttgccgaatgtgaacatgttatctaatcatgagtg
ctttccacgtgctagtgttgctagatgttatttcttcagcctaaaacaagctggggcctcagatgaccttcc
catgtagttcacagaattctgcagtggtcttggaaacctgcagccacgaaaagatagattacatatgttggag
ggagttggtaattcccaggaactctgtctctaagcagatgtgagaagcacctgtgagacgcaatcaagctgg
gcagctggcttgattgccttccctgcgacctcaaggaccttacagtggttagtatcaggaggggtcaggggc
tgtaaaagcaccagcgtagcctcagtggttccagcacgattcctcaaccattctaaccattccaaagggtat
tatcttggggggtgacattcttttctgttttcttttaactcttttttaaacatagaatttaataatatta
tgagcttttcagaagatttttaaaaggcagtcagaatcctactacctaacaacaaaaattgttttatctt
gaataatatgttcttgttgcattttccatgcagtgatgttaggcatacaaaatacattttttaagaa

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tactttcattgcaaattggaacttcgtttaaaaaatgctcataactaaaattggcattttctaaccatagggc
ccacttgtagttatattaccgaagcaaaaggacagctttgctttgtgtgggtctggtagggttcattagaaag
gaatgggggcggtgggagggttgggtgttctgttctctctgcagactgaatggagcatctagagttaagggtta
gggtcaacctgacttctgtacttctaaattttgtcctcag**GTCAATCCTGACCGGGTTGTTCCCCCGACC**
TCGGGCACCGCCTACATCCTGGGAAAAGACATTGCTCTGAGATGAGCACCATCCGGCAGAACCTGGGGGTC
TGTCCCCAGCATAACGTGCTGTTTGACATgtgagtaccagcagcagcttaagaataggccttttctggatgt
gtgtgtgtcatgccatcatgggaggagtgggacttaagcattttactttgctgtgtttttgttttttctttt
tttcttttttatttttttgagatggagtctcgtctctgtagccaggctggactgtagtggcgcatctcggt
cactgcaaccttggcctcccaggttcaagcgatttctcctgacctcagcctcccaggtagctgggactctagggc
acacaccacctgcccagctaatttttgtgttttttagtagagacggggttccaccatgttggccaggatgggt
ctcaatgtcttgacctgctgatccgcccacctcgggtctcccaaagtgtgggaacacaggcatgagccactg
tgtctggccacattttacttttctttgaatatggcaggctcacctccgtgaacaccttgagacctagttgttc
tttgatttttaggagaagtgggagggtgaatgggttgagctgtagaggtgacatcagcccagccagtggtgggg
gcttgggaaacattgcttcccattattgtcatgtctggaggggcctttagcccatcctctcccccgccaccc
tcttatttgaggcctggagcagacttcccagacctggtagtgcttcaggggccttggtatgatggacctatat
ttgtgtcttaagacatttgtctccactcaggttgtcccatcagccataaggccccaggaggacctgtgat
ggagcagagagagacctgagctctgcaatcttgggcaaggcttttcccttatgtttcttcttataaagtg
aacagctggggctcatgtgtctcctcctcatctaaagtgaacacatggggctcatgtgcagggtcctccccg
ctttcagagcctgagggtcccctgagggtcaggaaggctgtctccagggtgagtgcagagctgacttcttggtgg
acgtgtgtgtggggacagccattaaagaccacatcttggggcctgaaattgaaagttgtaactgcctgggtg
catggtggccaggcctgtctggaaacaggttgggaagcagctgtgtcacctttcactttgatttctgagcagct
catgtggttgtcactgttgttctaccttgaatcttgaagattattttcagaaattgataaagttatttta
aaaagcacggggagagaaaaatatgccattctcatctgttctgggccaggggacactgtattctggggtat
ccagtagggcccagagctgacctgcctccctgtcccccag**GCTGACTGTGAAGAACACATCTGGTTCTATGC**
CCGTTGAAAGGGCTCTCTGAGAAGCACGTGAAGCGGAGATGGAGCAGATGGCCCTGGATGTTGGTTTGGC
ATCAAGCAAGCTGAAAAGCAAAACAGCCAGCTGTCAggtgcggccagagctaccttccctatccctctcc
cctcctcctcgggtacacacatgcggaggaaaaatcagcactgccccagggtcccagggtgggtgcggttgg
taacagaaacttgtccctggctgtgcccctaggctcctctgccttactcactgtctggggctgggtcctggag
tttgtcttgtctgtttttttgtag**GTGGAATGCAGAGAAAGCTATCTGTGGCCTTGGCCTTTGTGCGGGGA**
TCTAAGGTTGTCAATTCTGGATGAACCCACAGCTGGTGTGGACCCTTACTCCCGCAGGGGAATATGGGAGCTG
CTGCTGAAATACCGACAAGgtgcctgatgtgtattttattctgagtaaatggactgagagagagcggggggct
tttgagaagtgtggctgtatctcatggctaggcttctgtgaagccatgggatactcttctgttakcacagaa
gagataaagggcattgagactgagattcctgagaggagatgtgtgtctttattcatctttttgtcccaac
atgggtgcactaaatttatgggttagttgaaaggggtggatgcttaaatgaatggaagcggagaggggcaggaag
acgattgggctctctgggttagagatctgatgtggtacagtatgaggagcacaggcaggcttggagccaactc
tggcztgggcctgagacattgggaaagtacaaacttgccctcaccttctttgccgataataatagtggtgczt
tacctcatagaggattaaatataatgagaatgcacacaaaccacctagcaccaatgcctggcatatagcaagt
tcccaataaaaatgcztactgttcttacctctgtgaggatgtggtacctatataacaaagctttgccattc
taggggtcatagccatacagggtgaaaggtggcttccagggtctcttccagtgcttaccctgctaatactc
tctagtccctgtcactgtgacaaatcagaactgagaggcctcacctgtcccacatcttgtgtttgtgacctg
gcag**CCCGCACCATATTCTCTCTACACACCACATGGATGAAGCGGACGTCCTGGGGACAGGATTGCCATC**
ATCTCCCATGGGAAGCTGTGCTGTGTGGGCTCCTCCCTGTTTTCTGAAGAACCAGCTGGGAACAGGCTACTAC
CTGACCTTGGTCAAGAAAGATGTGGAATCCTCCCTCAGTTTCTGCAGAAACAGTAGTAGCACTGTGTATAC
CTGAAAAAGgtgagctgcagctcttgggtgtctgggctgggtgttgggtctgggcagccaggacttgcctggctgtg
aatgatttctccatctccaccccttttgccatgttgaaaccacctctccctgctctgttgccctttgaaa
tcatatcatacttaaggcatggaaagctaaggggcccctctgctccattgtgctagttctgttgaaatcccg
tttctttttctatgaggcacagagagtgatggagaaggctccttagaggacattattatgtcaaagaaaaga
gacttgtcaagaggtaagagccttggctazcaaataagcctgggtzgttctgtctcattacttttcaatctcat
tgaccttaacttttaactataaaacagccaatatttattaggcactgatttcatgccagagacactctggg
caztgaaagaaagtaatgataatagttaattttatatagcgttggtaaccatttacaaccttttttttttt
taacctctatcatctcaattaaagtgcagagagaccttgggaagaaggtaactatatttattatcccagatg
agggaagtgaggcttgttagggaattggtagctgattcaagggtcacccagcaggtaataaacagtggtgggac
cagacccaattaccagggtatgttttctctgtaccgcagtacatgcctgagatttatttgtgtgttgaaagcc
agtggtagcctaattgtatttacctcccaacctgaaactcctatccacttatttaccttttaatgagcctctta
actcaagtgcagtctgaggaccagcagcatcaggatcacttgggaacttgttagaaattcagcaacctgggc
ccagctcagacctaccgaatcagaatctgtgcattttaacaagggtccttgagtgggtgaaacacacattaaag
catgagaagcattgaactagacatgttagccagggtaaaggccttgccctgagatgggttggcaaggccctcattg

bioRxiv preprint doi: <https://doi.org/10.1101/123456>; this version posted January 1, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cagcattcattggcagggccacagttcttttggcagctctgcttctgacctttcacccctcaggaagcgagggc
tggtcacacggccacacacatgccagacaggggtcctctgaagccacggctgccagtgcagtggtcccagggaa
agcttttttcttttagttctcacacaacagagcttcttggaaagccctccccggcgaaggtgctgggtggctctg
ccttgctccgctccctgacccgttctcacctccttctttgccatcag**GAGGACAGTGTCTCTCAGAGCAGTTC**
TGATGCTGGCCTGGGCAGCGACCATGAGAGTGACACGCTGACCATCGgtaaggactctgggggtttcttattc
aggtgggtgacctgagcttccccagctgggcagagtgaggagcagaggagagaggtgcagaggctgggtggcgc
tgactcaaggtttgctgctgggctggggctgggtggctgcgggkgtgggagcagcttgggtggcgggttggcc
taatgcttgctggggtgacctggggctcgggttgggagctagcagggcagtgctcccagagagctgagatgatt
gggggttggggaatcccttaggggagtggaactgaataaccagggatgaggagctgagggccaagccaggag
gggtgggatttgagcttagtacataagaagagtgaagagccaggagatgaggaacagccttccagattttctt
tgggtagcgtgtgtaggaggccagtgctcaccagtagcatatgtggaacagaagcttggaccttgctatctc
tgacctagtccctaatggctggcttttccaggaaggtcttgccttccatggactgttagattaaccctttatt
taggtaaatgagggaaacctactttataagcataggaaaggtgaagaatcttttaagattcctttactcaag
ttttcttttgagaatcccagagcttaggcaatagacaccagactttgagcctcagttatccattcacccat
ccaccacccacccacccatccttccatcctccatcctccattcacccatccaccatccagctgtccac
ccattctacactgagtacctataatgtgacctggcttgggtgatacaaaggtgaataagacatagctcctttcc
tttggccccaacccctcagaccagagatgaacatgtggaatgacctaaacacctggaacaggtgtgggtgatg
agcggcagggcctctgatgagaggggtgggggatggccagccctcactccgaagccctctgagttgattgagc
catctttgcattctgggtccctgcag**ATGTCTCTGCTATCTCCAACCTCATCAGGAAGCATGTGTCTGAAGCC**
CGGCTGGTGGAAGACATAGGGCATGAGCTGACCTATGTGCTGCCATATGAAGCTGCTAAGGAGGGAGCCTTT
GTGGAACCTCTTTCATGAGATTGATGACCGGCTCTCAGACCTGGGCATTTCTAGTTATGGCATCTCAGAGACG
ACCCTGGAAGAAgtaagtttaagtggtgactgtcggaatatatagcaaggccaaatgtcctaaggccagacc
agtagcctgcattgggagcaggattatcatggagtttagtcattgagtttttaggtcatcgacatctgattaa
tggtggccccagtgagccatttaagatggttagtgaggagatagcaggaaagaagtggttttctctgtaccaca
gtacatgcctgagatttgtgtgttgaaaccagtggtacctaacacatttacatcccaaccttaaactcctat
gcacttattttaccccttaatgagcctctttacttaagtacagtgkgaggaacagcggcatcaggatcacttg
ggaacttggttagaaattcagcaacttgggcccagctcagacctactgaatcagaatcaggacaattctctg
gtgtgactgtgtcacagccaggtatcaactggattctcatacataggaaatgacaaacggttatggatggat
agtctacttggtgccaggtgctgagatttgtttttgttttttgatttttttttaactactgtgacctcattt
aattctcaaaaaaagatgaaaaaatgaacactcaggaatgctgacatgagattcagaatcaggggttggggg
cttcaaagttccatcctctctttatccatgtaatgcctcccttagagatacaacatcacagacctgaaggc
tgaaggggatataaaaagctgtctggccaagtggtctccaagcttgacagtgcagcagaatcacctggggata
ttattaaaaataaacataactaaggttgggttcagggcctgtgaatcagaatttctggaggtgaggccttga
agtctgtattttctattgcatactttggacacagtggtctatagactagagtttggaaatgattgcgctcatt
cagattctctctctgatgtttgaattgctgccatcatatttctagtgtctatttccctcctgctcattctgtc
ttggataacttatcatagtactagcctactcaaagatttagagccacagtcctgaaagaagccacttgactc
attccctgtaggttcagaataaatttctctgctgcagtgctgtcatagcttttttttaatttttttttatt
tttgatgagactggagttttgctcttattgcccagctggagtgagtgagtggtgagatttgggtcactgcaac
ctccacctcccaggttcaagcgattctcctgcctcagcctcccaagtagctgagattacaagcatgtgtac
cacgcccagctaattttgtatttttagtagagatgggttttatccatgttgggtcaggctgggtctcgagctcc
agacctcaggtgatctgcccgcctcggcctcccaaagtgtgggattataggcctgagccacagcgctcagc
cataactttaatttgaaaatgattgtctagcttgatagctctcaccactgaggaatgttctctggcaaaa
cggcttctctcccaggttaactctgagaaagtgttattaagaaatgtggcttctacttctctgtcttacggg
gctaacatgccactcagtaataataaatcgtggcagtggtgactactctcgtaatgttgggtgcttataatg
ttctcatctctctcattttccag**ATATTCTCAAGGTGGCCGAAGAGAGTGGGGTGGATGCTGAGACCTCAG**
gtaactgccttgaggggagaatggcacacttaagatagtgccttctgctggcttctcagtgccagagtattg
ttcctttccctttgaattgttctattgcattctcattttagagtgtaggtttgttgcagatggggaaggtt
tggtttgttgtaataaaataaagtatgggattctttccttgtgccttcag**ATGGTACCTTGCCAGCAAGAC**
GAAACAGGCGGGCCTTCGGGGACAAGCAGAGCTGTCTTCGCCCCTTCACTGAAGATGATGCTGCTGATCCAA
ATGATTCTGACATAGACCCAGgtctgttagggcaagatcaaacagtgctcctactgtttgaatgtgaaattct
ctctcatgctctcacctgttttcttggatggccttagccaaggtgatagatccctacagagtcctaaagag
aagtgagggaatggtaaaagccacttgttcttgcagcatcgtgcatgtgatcaaacctgaaagagcctatc
catatcacttcccttaaaagacataaagatgggtgcctcaatcctctgaacctatgtatttattatctttctg
cggggctcctagtttcttgtatacattaggtgtttaattgttgaaacaaatattcattcgagttagatgagtgat
tttgaaagagtcagaaaggggaatttgtgttagagtttaattgtaccctaagacttagatatttgaggctgg
gcatgggtggctcatgccagtaatcccagcgctttagagaggtgaggtgggttagatcacctgaggtcaggagt
ttgagaccagtctgaccaacaaggtgaaaccccgctctctactaaatacaaaaaattagccgagtggtggg

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

acatgcctgtcatcccagctacttgggagggctgagggcaggagaatcgcttgaacccaggaggcagaggttgc
agtcagccacgggttgcgccattgcactccagactgggcaacaagagtgaactccatctcaaaaaagaaaa
aaaaagaattagatattttggatgagtggtgtctttgtgtgtttaactgagatggagaggagagctaagacat
caacaaaatattgttaagatgtaaaagcacatcagtttaggtatcattagtttaggacaaggatttctagaaa
atTTTTtaggaacagaaaactttccagtttctctcaccctgctcaaagagtgtatggctcttacattatata
aactgcctgacttcatacagtatcagtttagatcatttgaaatgtgtccacgttttaccaaaatataata
gggtgagaagctgagatgctaattgccattgtgtattctcaaatatgtcaagctacgtacatggcctgtttc
atagagttagtctataaagaattgatgacttgattcatccgaatggctggctgttaacacctggttacgcatga
acacctcttttccagttgtctcaagacacctttcttttctgtacttatcagacaaggactgaaaggcagagac
tgcctactgttagacattttgagtcagcttttcccttggacatagcttgtcatgaaagcccttacttctga
gaaacttctagcttcagacacatgccttcaagatagttgttgaagacaccagaagaaggagcatggcaatgc
cgaaaacacctaagataataggtgaccttcagtggttggcttcttgcag**AATCCAGAGAGACAGACTTGCTCA**
GTGGGATGGATGGCAAAGGGTCTTACCAGGTGAAAGGCTGGAACTTACACAGCAACAGTTTGTGGCCCTTT
TGTGGAAGAGACTGCTAATTGCCAGACGGAGTCGGAAGGATTTTTGCTCAGgtgagacgtgctgttttgc
ccagagactctggcttcatgggtgggtgcaggtctgtgaccagtgaaggcaggatagcatcctggtcaag
atatggatgccggagccagatttatctgtatttcaatccagttctattccttggcagttgtgtatccgctg
gcaagttacttctctatgcctcaatctcctcatctgtaaaatggggataataatattacctgcaatacaggg
ttgttacgaaaataaaaaatgaataggtgcttagaatggggcctgacattagtaagtgttagttttgtgtgt
gtatatgttatttttattttggaggagaacataaaaaggacaaagtgtagaaaaactggttgggtgtattca
gctgtcataacatgagagttgttatgccagatgcacttgacatgtgaatttattagaaacatgatttttct
ctgagttgatgtttaactcaaactgatagaaaagataggtcagaatatagttggccaacagagaagacttgt
tagactattgtctgcatgtcagtggttgcatgctaacttgcttagttagaaaggttaaattttttcactcta
taaaatcaagaaatatagagaaaaggtctgcagagagtcctttcatttgatgatgtggatattgttaagagcg
ggagtttggagcatacagagctcaagttgaatcctgactttgctacttattggctatatgaccttgggcaag
ctgcttagtctctctgatcctcagttacctttgtttgttgatgatgaccattgataacacaaccataataa
tgacaacatagagatagttctcattatagtagttgtttatacagaattattcactcaatgtttaattttctgca
ttgaaatcccagaacattagaattggggggcattttgaatctttaagggtataaggaatacatttctcagc
aataaatggaaggagtttgggttaacttataaagtatacccaagtcattttttttcagagaagatatgggt
agaaagtcttaggaggttgaagaaggaattggatatttattctttctgagactatcatgggagataatgact
atgggtgtccatgattggagccgttgcctgtagagttgggttttattatagtgtaggatttgaatgggccatgt
gttctcagacctcagattaaaawgagaaaactgaggccagtggggagcgtgacttcacatgggtacacttgt
gctagagacagaaccaggattcaggacttctggctcctggctcctgggttcatggcccaatgtagtctttctc
agtcttcaggaggaggaagggcaggaccagtggtctgagtcacccctgaatgtgagcactatttacttcgtg
aacttcttggcttagtgccctctgccaggtggccataacctctggccttgtgttgccagagaaaagggttagt
tttcagggtccattgtctcccagctgccaagaatgccttgggtgcagcacagtcataggccctgcattcctca
ttgccgtgctgggttgggtcggggaggtgggctggactcgtagggatttggcccttggccttgtttctaacact
tgccgtttcctgctgtccccctgccccctccactgcctgggtaaag**ATTGTCCTTGCCAGCTGTGTTTGTCTG**
CATTGCCCTTGTGTTTCAGCCTGATCGTGCCACCCTTTGGCAAGTACCCAGCCTGGAACCTTCAGCCCTGGAT
GTACAACGAACAGTACACATTTGTCAGgtatgtttgtcttctacatcccaggagggggtaagattcgagcag
accaaagatgtttacgagggccaagggaatggacttcagaattacacggtggaatgaattttactgctgcgg
ctcaggtccctgtataagctaatactgcagtcagtagaacagcagcgaactaacctgaataataggccagtc
ttctgttgagcctttcagcctctctcctctcactactgttgtcaggaacagccacatgtgttttaggtg
aaataatccaccttgcaaaaaatccatgattaagttataaaaatatttggatttgtggagctgtgttttaatt
ctgtaactgagtcacagggcacactgtcaaagcatagaacctccagagacttgttttctgcaaaagtataatt
catgtaattattatctattctgttatatttgggatgttaggtagtggtttgttcttttagataaaaaatataccc
cactctgtaacaatacattaaatcaaagaaaaggacaaaggatttttctgggtcttgttagcaggagctttc
ttcagtcctgaaagatttgtagacctgtagatgggggaactgtgtcagtgatacaaaagggaagcattttaa
aaaaaaaaagtatatatatatatatatatatatatgtaatgtgaattggcctctttttctctaagzcca
cattttzcttcttacatagttcagggttactttattttttcctttccggctgctgacctgtattgcccgta
gttgtggaacatagcatgtgtttgtgacctgtgcctgttatttttgtgtcttctagttgtgcatgcaaagag
tacaaagttttcttggcccttcttggaaaatcctgcttgtctgtgccaaggataattgtgaaagcacttt
tgaaatacttaatgagttgattttcttcaaatataaaaaaatatataaatgtatatgtgtatgtacatgtgt
gtacacatacacacctttatacatacagcccatttaaaacaagctccactttggagtgctctacgtcacct
gatgccgaatacagggccagagtcctgagatccttctgggtgggtttctgtgttttgttcatttctgttttaag
agcctgtcacagagaaatgcttccataaatgttttaattataaaaaacatttttatctctcgattacttggttt
taatgaattactaagctggctgcctctcatgtaccacag**CAATGATGCTCCTGAGGACACGGGAACCCCTGG**
AACTCTTAAACGCCCTCACCAAAGACCCTGGCTTCGGGACCCGCTGTATGGAAGGAAACCCAATCCCgtgag

GenBank accession number: F01430.1

cggggtgggtcacacctgtaatctcagtatcttgggatgccgaggtggactgatcacctaaggtcaggagtt
cgaaaccagcttggccaatgtggcgaaacctgtctctactaaaaattcaaaaattagccaggcatgggtggc
acatgcctgtaatcccagttacttgggaggtgaagcaggagaattgcttgaacctgggaggtggaggtcgc
aggagaccaaaattgcccactgtactccagcctgagcaacacagcaaaactccatatcaaaaaataaaatg
aataaaataacagctaatactagtcacagtaataactccagtgaaacagaagatttattaggtatagtgaaatga
tgggtgcttccataaaaatctcttgactacaaagaatctcatttcaatgtttattggttagatgttcagaataa
attcttgggaagaccttggcttgggtgaagtgaattaccagtgccgagggcaggggtgaaccaagtctcagt
gctgggtgactgagggcagtgcttgggacctgtagtccaggttccgggtcacactgtggacatgggtcactgtt
gtccttgatttgttttctgtttcaattctgtctataaagaccgtatgcttgggttttcatgatgatgacagA
GAAACAAAACACTGCAGATATCCTTCAGGACCTGACAGGAAGAAACATTTCGGATTATCTGGTGAAGACGT
ATGTGCAGATCATAGCCAAAAGgtgactttttactaaacttggccctgccktattattactaattagagga
attaaagacctacaaataacagactgaaacagtggggggaaatgccagattatggcctgattctgtctattgg
aagtttaggatattatcccaaactagaaaagatgacgagagggactgtgaacattcagttgtcagcttcaag
gctgaggcagcctgggtctagaatgaaaatagaaatggattcaacgtcaaattttgccacttagtagcaactt
gaccaggtaactgggttatccttttaagccttagtttatctaaattgtgatattaatgttgcctctataagt
ttgtcatgaggactaaattaaatgggtgtacatagagtgccttgggtactctctgatgggggactccatgata
atttgtgggtctcatggaggagctctgggaaggtttaggagcctgccttgggtctgcagccttgggagagcc
ttctagcttcccaggacatggcagcctagtgttgaatgcttgggtcagcaaatgtttgttctcgtttccctc
ccatcaacttgggtcagttgggggtcttccagttaggagtatctcagtgactttaaatggcatgggcatgctgg
agtgatagtgaacatgagtttctaaagaagaagcataatttctccatatgtcatccacaattgaaatattat
tgttaattgaaaaagcttctagggcaggcacggtgggtcatgcctgtaatcccagcactttaggaggccaag
gcgggtggatcacttgaggtcaggagtttgagaccagcctggccaacatggggaaacctgtctctactaaa
aatacaaaaataagctgggcgtgggtgggtgcgtgcctgtaatcccagctacttgggaggtgagggcaggagaat
tgcttgaatctgggagggcggaggttgacgtgagctgagttcatgccattgcattccagcctgggcaacaaga
gcgaaccatctcccaaaagaaaaaaaagaaagaaaagcttctagtttgggttacatcttgggtctataag
gtgggttgtaaaattgggttaacccaaggcctgggttctcatataagtaataagggtatttatggagagaaag
gctggaagaggcctgaacacaggtcttcttctagcacaccctacaaggccagctgattctagggttat
ttctgtccgttccctatatccctcaggtggatatttactccttttgcacatttaggaataggctcagtgcttt
ctttgaactgattttttgtttcttctgtctctgcagCTTAAAGAACAAGATCTGGGTGAATGAGTTTAGgttaa
gttgcctgtctttctggcacgttttagctcagggggaggatgggtgttgtagggtgtzcttggattgaagaaagcc
ttggggattgtttgtcactcacacacttgggtgccatctcactgtgaggaggacagaagccctgtgaaca
tgtggagcacacaggggacagacagatttagattaggcctgctttatagagtttctgcctagagcatcatg
gctcagtgcccagcagccctccagaggcctctgaaatatttgatatactgatttcccttaggagagaatcaga
aatctcctgcaggtgtctagggatttcaagtaagtagtggttgtgaggggaatacctacttgtactttcccc
caaaccagattcccagaggttcttaaggactcaaggacaatttctaggcatttagcacgggactaaaaaggt
cttagaggaaataagaagcgccaaaacctctcttgcactgtatttcaacctatttgccttctgggtttt
gaaggaacaggtgggactggggacagaagagttcttgaagccagtttgtccatcatggaaaatgagatagg
gatgtgggtacgtcagggggcccgaggctccttgttactgatttccgtcttttctctctgccttttcccc
agggccaggacccctggatctctgggcagagcagacgcagggccctataatagccctcatgctagaarggag
ccggagcctgtgtataaggccagcgcagcctactctggacagtgcaggggttccactctcccaactcccc
ctgcttgcctccagaccacattcacacmcgagccactgggttggaggagcatctgtgagatgaaacacccat
tcttccctcaatgtctcagctatctaactgtgtgtgaatcaggccaggtcctccctgctgggcagaaacca
tgggagtttaagagatttgccaacatttattagagggaagctgacgtgtaacttctzzgaggcaaaatttagccc
tctttgaacaggaatttgactcagtgaaacctgtacacactcgactgagtcgtctgctgatgatactgtg
caccctactgtctgggttttaatgtcaggtgttcttttagGTATGGCGGCTTTTCCCTGGGTGTCAGTAAT
ACTCAAGCACTTCCTCCGAGTCAAGAAGTTAATGATGCCATCAAACAAATGAAGAAACACCTAAAGCTGGCC
AAGgtaaaaatctctatcgtaagatgtatcagaaaaatgggcatgtagctgctgggatataggagtagttggc
aggttaaaccgatcacctggcagctcattgttctgaatatgttggcatacagagccgtcttggcatttagc
gatttgagccagacaaaactgaattacttagttgtacgtttaaagtgtagggtcaaaaacaaatccagagggc
caggagctgtgggtcatgcctgtaatcctagcacttgggagggctgaagcgggtggatcacttgaggtcagg
agttcgagaccagcctggcctacatgacaaaaccccgatctactaaaaatacaaaaaaattagctgggctt
gggtggcacacacctgtaatcccagctacttgggaggtgagggcaggagaattgcttgaacctgttaggaaga
gggtgttagtgagccaagatcgaccgttgcactccagcctgggcaacaagagcaaaactccatctcaaaaa
caaattaaatccagagattttaaagctctcagaggtgggcgcgggtggcttacacctgttatccagcattt
tgggatgccgagggcgggcaagcacaaggtcaggagtttgagaccagcctggccaacatagtgaacacctgt
ctctgctaaaaacatagaaaaatttagccgggcatgggtggcgtgcgcctgtaatccagctactcgggaggt
gaggtgagagaatttcttgaacccgggagggcggaggttgacgtgagccagattgcaccactgcactccagc

[illegible]

acccatctcttgtcttctcctgggtattatctgtccctccctgcttttagagctcctgaaatttgctagaagcat
 gtcttcatctaaagttgttgataaacacatcaagtaggattggactgaggcagagccctgtagctgaagctg
 cagttcttctagcggctgacaagccccactatcacttccctgctgggtgctttgctctgccagctgtgaattc
 tcataattgtcctatcgtaagtcctttatcttctgcattttactgcttgatacactgtcaggacagactttaa
 aattattctcagtgccgatgaaacaattctgacattcatgttatgagcagttacctcataaatagattacatg
 tgagattgaacttgggcagactataatatagcattaatgatgaaacagacacagtcattctcggaagaaga
 atagagggttatttgctgctgtgaaattaaaattactctgactgggaatccatcgttcagtaagtttactg
 agtgtgacaccttggcttgactgttggaagacagaaagggcatgtagtttataaaatcagccaaggggaaa
 atgcttgtcaaaatgtattgtcgggtattttgattaatagtttatgtggcttcattaattcagagttactct
 ccaatatgtttatctgcccccttctgtctgataatggtgaaaacttgtgtgatgcattgtatatttgattta
 ggggtgaactggatgtctttgttttcttctttag**TGCAATTACGTTGTCCCTGCCACACTGGTCATTATCAT**
CTTCATCTGCTTCCAGCAGAAGTCTATGTGTCTCCACCAATCTGCCTGTGCTAGCCCTTCTACTTTTGTCT
GTATGGGtaagtcacctctgagtgaggagctgcacagtgaggaatggcatttggtgcccagtggtcagaagga
 gggcagggactctcagtagacacttatcttttgtgtctcaacag**GTGGTCAATCACACCTCTCATGTACCC**
AGCCTCCTTTGTGTTCAAGATCCCCAGCACAGCCTATGTGGTGCTCACCAGCGTGAACCTCTTCATTGGCAT
TAATGGCAGCGTGGCCACCTTTGTGCTGGAGCTGTTTACCGACAATgtgagtcagtcagagagaacactcct
 gctgggatgagcatctctgggagccagaggacagtggttaattgtgatcttattccacttgtcagtggtatt
 gacactgctgactgcttctgtctcagagtcgtcttccctgagaaggcaaagcaccttcttctcttg
 ctgtgcttacattttgtgggtcaagccttccagtttcttttgacagtttttttacttcttcttttttca
 atgttgccttaccagagtagctcctctgcttccactttacacatgagagctgggcgacgzcattcagtc
 ctaaggcttttaccatcacctctcttgggtgtttttattgtcatctctaagatcaatgcctttagccttgatc
 ataacctgaactctaattctcaaattctcacttgcctagtggattgctccatttagatagatatagatacc
 ccaacctggatatgtccttagttttcttctcccttggaaacttaatgcttttcttgccatccctgtcacactca
 gtggcactaccatccactcggttggccaaagctggctcttagagttatccttagatgcttgctttgctgttgca
 gatttcccacattcaactgggtatgttgtcagttcttccaggtatggacctctaaaataaggcttccctctcc
 attccggttgtcattgccttctgccaacacagcacacaaggccttttacagttgcacaactcttccctgtcc
 ataccacacacaccttctccagctgtaagcttcagatgagttgcctccaaccacctgctcctgtaggcct
 ggcttgaaattgcccttctctgtcacaggggtcgtgttagtatatcccttgccctcaagatttagctaaaatg
 tgaagcttcttacctgctgggaggtgtctctctcttctctctgtgctctcagagtccttagtccatgcctc
 cagtacaaactacatccacttacatggtaatttctctgtttacatacttcttctactcggagtgaggtctgtt
 tcttaataattttgcctctcccatgccctagcacagtgcatccagcgtatagcccccttattcagttggtaga
 tatttggccactgttgcttgtgggatcataagttctgatgtatttgagaagaatttctaaaattctgacaa
 aatcctgaaactcaaatattgaccacagacatgagcaatttgcttttcaaatgctaagggtttttaaaggat
 ttgctttaattaaatctagcctgtttctaagctttattcattatttctccatactcagagcatttctccaga
 ttttctaaagaatagaattttattgctacatatcatcagctatgcctgctgctatttaattgggtatctgaat
 taaaaggctcgtgttgccttagagaatcaaatttttcttccactcccatatttccagaacttgatacatttt
 taggataaaccatgaatgacacccgtttcttctccctcaccctcccttccctccctttttttttttttttt
 tttttag**AAGCTGAATAATATCAATGATATCCTGAAGTCCGTGTTCTTGATCTTCCCACATTTTGCCTGGG**
ACGAGGGCTCATCGACATGGTGAAAAACCAGGCAATGGCTGATGCCCTGGAAAGGTTTgtgagtgaaagcag
 tggctgtaggatgctttaatggagatggcactctgcataggccttggtacctgaactttggttttggaaga
 agcaggtgactaagcacaggatgttccccacccccatgccagtgacagggctcatgccaacacagctgggt
 tgtggcatgggttttgtgacacaaccatttgtctgtgtctctgatagcattgagaaaagtgaagggcagtt
 ttgaaggtaaggaaaatagtgttatttgcctggatccactggctcatgccactgtctgggttggttagaagc
 actggaaaagtcaaaccataactttgagaattaggtgatcaggggaatcagaaggaaagatgcaaaccttggc
 tcttttaggcgaatcatgtgcctgcagatgaggtcatttattatcttttacacagtcataaaaattataatg
 tattacatcttttctacctttagaatgggttaaaaaatatttctccggtagccatattgattatttcatcca
 ttagataatatagtcaaatgggcatgttatttactgttcatagaagaggggctttttgcaacttgggctac
 aaaggagatatgtaagggaatttaagggaatgggtacatggaactagatttaattgaatctagtgggttaattg
 attcactaggatatatgctactgaaaggggaatctgcttaaagtgccttctgatatttatttactaaaac
 ttagaatttattaaaaaactgactgtgaaaattacttgggtcggttgccctttttaaaggatttttggcat
 gtctcattaaaaaaaagaaatactagatatcttcagtgaagttacaaatcgaatacacattggctctgaaatt
 ctgattgatactgggtcataaaaagttttcccaaatcagacttggaaagtgatcactctctgttactcttt
 tttccttgtcatgggtgatagccatttgtgtttatttgaagatcggtgaattttaagggaacataggcccaa
 tttgaggaagggtcatggtttttgatccctccattctgaccgatctctgcattgtgtctactag**GGGAGAA**
TCGCTTTGTGTACCATTATCTTGGGACTTGGTGGGACGAAACCTCTTCGCATGGCCGTGGAAGGGGTGGT
GTTCTTCCTCATTACTGTTCTGATCCAGTACAGATTCTTCATCAGGCCAGgtgagctttttcttagaacc
 gtggagcacctgggtgaggtcacagaggagcgacagggaaacactaccaatgggggttgcatgaact

gaaactcaaaatatgtgataaaaactgatttttctgtagtggtggcgcaccccgagcccccctcctgcccattcctgg
agactgtggcaagtagggttttataactacgttagagctgaatctttgtcctgaaaaatagtttgaaagg
ttcattttttctgtgtttttcccccaag**ACCTGTAAATGCAAAGCTATCTCCTCTGAATGATGAAGATGAAGA**
TGTGAGGCGGGAAAAGACAGAGAATTCTTGATGGTGGAGGCCAGAATGACATCTTAGAAATCAAGGAGTTGAC
GAAGgtgagagagtacaggttacaatagctcatcttcagtttttttcagctttatgtgctgtaacccagcag
tttgctgacttgcttaataaaagggcagtggttcccaaaatgtacatctataccaagggtctgtcaatttta
ttttaaaaacacccatggagacttcttaagaattcttactgagaattcttttgtgatatgaattcccattct
cgaatacattggttttatatgcttacatttatgtgttagttattaaaacataactaatattgtatatctagtc
aaaactgaggttagagagaataaatggttgattttgagtttgagtttcatagtccaaaaagctgatataattgc
ctgtgttcaagaggggtctatatcagccctctagatgccagcatctccaaattttacttttttggaaactgt
cagtatttgcaatatattttattacaaatttctactctgtggaatttaatttttaaaatacctgcaatacata
tatatgttgtaatatagatgaaaaattatgtagatrataatgaatgatacgggtctaaaaagacaggttaaaaag
taagttcacttttattttgagcttcagaatcattcagaagccagtcgccacaaacgcagacgaaggctcttg
gcacatcaaatatgcctatggccttagggtattgacaagcttatgttgcagtgtagtggtttatagtcct
gccttccacagttgcttggggagagctgtgagtcactgaggccttatgaatgtttacattttgtttgttgca**A**
TATATAGAAGGAAGCGGAAGCCTGCTGTTGACAGGATTTGCGTGGGCATTCTCCTGGTGAGgtaaaagacac
tttgtctatattgcgtttgtccctatttagttcagactatctctaccaatcaagcaacgatgctcgttaa
ggtaaaagtggattttaaaggcctctgtatttatgccaggatggagcaatttagtcatcgagaagagagggac
cctgtatgtcaagagaatgatttcagagaatccaatacaatttaagaaaaagcatggggctgggcgcagtgat
ttcactcctgtaatcccagcactttggggaggccgaggtgggcggactcacgaggtcaggagattgagaccat
cctggccaacatggtgaaaccccatctctactataaaatacaaaaatttagctgggcatagtagtgcatcctctg
tagtcccagctactcgggaggctgagggcaggagaattgcttgaacctaggagggggaggttgcccagattgc
gctgctgcactccagcctggtgacagagtgagactcatgtcaacaacaaaaacagaaaaagcacgcacatct
aaaacatgcttttgtgatccatttgggatgggtgatgacattcaaatagtttttaaaaatagattttctcct
ttctggtttccgtttgtgttcttttatgcccttttgccagagtaggtgggtgczaatttggctazgctggctt
tcattactgtttttcaczacattaaacztttggcctcaacttgacaactcaaaataattataaatacagc
cacacttaaaatgggtcccattatgaaatacatattttaaatactatacagatgtgttaaaacagaaaaaat
ttgattcttctctgataatttaagaattgaaggtttagggtagttacgtgttaggggcatttatattcatgtt
tttagagtttgcctatacaacttaacttttcttcttcag**TGCTTTGGGGCTCCTGGGAGTTAATGGGGCTGGA**
AAATCATCAACTTTCAAGATGTTAACAGGAGATACCCTGTTACCAGAGGAGATGCTTTTCCTTAACAAAAAT
AGgtgagaaaaagaagtggccttgatttttgcctgcaaagactttgttttttaattttatttaagaaataggttgt
tatttttgattacagtggtatttttagagttcataaaaaatgttgaaatatagtaaaagggtaaagaagcacat
aaaatcatccatgatttcaatatctagagataatcacaatttacatttcttctcagtcctattctcttctt
taacagctttattcaggtataatttacatacaatataatttgcttggtttttaagagtataatttagtgatt
tttggtaaatgagagttttgcaaccatcaccacaatccagttttagaacttttccatcaccccacatctgt
cttatatacacatataaatgtgccatacaattgagatcatactgtatgtagaatttaaaattagttttatt
gttaatgagtgattatgaatatattcccagtggggttacatttccctaagatgtggaattttacattgctacat
aaaaatccccctatgtacatgtacctataatttatttaataaaattccttataaatgttggacacattagtttc
catttttactatgtaaatatgtccctgtatacatcttttattatttctcagggaacaattctcacaaagt
aattgccctctctaaagagcatacaaattgactgagccaccggttaggccattttctgagactgcacaggtca
caaagcaatctgatcttctgggaatacagctacattttataggtctcttagaataatgttactctaaagtatt
aaatgctgagggtcttctgggccttttttttttgagacggaggttctactcttagtcccaggtcggagag
caatggcgcgccttggtcactgtgcaacctccgcctcccaggttcaagcgatttctcctgcctcagcctcctg
agtactgagattacaggtgcccgccacaatgcctgcctaatttttttgtattttcagtagagatgggggtt
caccatgttggccagactggctctcgagctcctgacctcaggtgatccacctgcctcagcctcccaaagttct
gggattacaggcatgagccactgcgcccgccttctctggacttattatgtggagagatagtacaaggcagtg
gctttcagagttttttgaccatgaccgttggtgggaatacattttatatctcaacctagtatgtacacacag
acatgtagacacatgtataacctaaagtttcataaagcagtagcctactgttactaattgtagtgcactctgc
tatttcttattctaccttatactgcgtcattaaaaaagtgtgtggtcatgaccactaaatttatttcccaa
ccactaatgaacaatgactcacaatttgaacacactggacagggggatagccaataaaattgaaaagagcaa
ggaaattaatgtattcatgatctcctctcctgtctcttacatttttgcagtagcaatgtaaaggaatcctaa
gagaacagacattctgggaatagcaggcctagcgtgcacaactgctttcctaggcttgctcctagtaccaa
gctcctgacgcataatagcagtggcagtaataaccagcccatagtaagggttggtcacagggaactgggtgtaag
aactgatttggttggatatagctgtgagggcctggcacgggtgtccagtggtgcctcaatcctaattctgaaaa
aggctgacctgggggtgctaatttagatacacagaggaatgaatgtgcagaaggccaaagttcatggca
atgcccgtctggctgaggtgcagtcagtcagtcggaaactggaactggaactctctcacatgtgattcttc
acttgactggcttcatagaaccccaagccacccaccaccacataaattgtgtctctaggttctgtgttg

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tcaactcaaaatttctggtccttctcatttgggtgcatgtgaatgggtgcatatgagtgaagtctaggatggg
gccttagcggttaaagccctggggtagtgtagtgcagattggttggtaaagaatgtgcagtggttggcatgacc
tcagaaattctgaaatgggactgcacctgcagactgaagtgttcagagagccagggaggtgcaaggactggg
gagggtagaggcaggaaccctgcctgccaggaagagctagcatcctgggggcagaaaggctgtgctttcaag
tagcagcagatgtattggtatcttgtaatggagaagcatactttacaggaacattaggccagattgtctaa
ccagagtatctctacctgcttaaaatctaagtagttttcttgcagTATCTTATCAAACATCCATG
AAGTACATCAGAACATGGGCTA/TGCCCTCAGTTTGATGCCATCACAGAGCTGTTGACTGGGAGAGAACACG
TGGAGTTCCTTGGCCTTTTGAGAGGAGTCCCAGAGAAAGAAGTTGGCAAGgtactgtgggcacctgaaagcc
agcctgtctccttggcatcctgacaatatataccttatggctttccacacgcattgacttcaggctgttt
ttcctcatgaatgcagcagcacaaatgctgggtcttctgtatctgctttcaggggtggaaacctgtaacgggtg
gtggggcagggctgggtgggcagagagggagtgctgctcccaccacacgagtccttctcctgctttggct
cctcaccagttgtcaggttatgattatagaatctagtcctactcagtgaaagaactttcatacatgtatgtg
taggcacagcatgataaaattcccaagccagaccaaagtcaagggtgctttttatcactgtagGTTGGTGAGTG
GGCGATTTCGAAACTGGGCCTCGTGAAGTATGGAGAAAAATATGCTGGTAACTATAGTGGAGGCAACAAACG
CAAGCTCTCTACAGCCATGGCTTTGATCGGCGGGCCTCCTGTGGTGTTTCTGTgtgagataactgtggatgg
aaaactgttgttctggcctgagtggaacatgactgttcaaaagtcctatatgtccagggctgttgtatga
ttggcttgtctccccagggacagcagagcaaccttggaaaagcagaggggaagcttctccttggcacaca
ctgggggtggctgtaccatgcctgcagatgctcccaatagaggcactccaagcactttgtttcttagcgtga
ttgagggtggatattgtgatttgatcttctctggaacattcttctaatcatctttgtgttcattccctgaa
aatgaagagtgtggacacagctttaaaatccccaggttagcaactaggtcatagttcctztacacacggata
gatgaaaaacagatcagactgggaagtggcccttgacctttttcttctgtagataagagcattgatgttat
tacgggaagaagcctttgaggttttatgtattccacctcgggtctggaatttggtttctgtaaggctaacagt
tgcaatatactagggtaatctgagtgagctggaattaaaaaaaaaaggaatttcaccccaatcttatactg
acttcaatagaggtttcagacaaaaagttgttttgatatatacttatcagtcagtaaaagataattacaacta
aatggcctttttctcctcctatttatttgggagaatttaattacataaaaaagtaactcagaatatttgagtt
tctgcatcaataagacatttataataatgacctgtttacaaatgaatttgaaagttactctaattctttg
attcatcaagaaataactagaatggcaagttaaaatttaagctgtttcaaagatgcttctgcatttaaaaac
aaatttatctttgattttttttccccccagcaataagacttattttatttctaattacagGATGAACCCACC
ACAGGCATGGATCCCAAAGCCCGCGGTTCTTGTGGAATTGTGCCCTAAGTGTGTCAAGGAGGGGAGATCA
GTAGTGCTTACATCTCATAGgtccgtagtaaagtccttgggttctcactgtgggatgttttaactttccaag
tagaatatgcgatcattttgtaaaaattagaaaatacagaaaagcaaaagagtaaaacaattattacctgaaa
ttatatatgcataattcttcaaaaaatgcaagcccagatataaatactgctctttttcacttaatatattgtaa
acattattccaagtcagtgcattttaggtgtcatttcttatagctggatagattccattaggatatactctt
atttaactattccccctttttagacatttggattatttccaacttggttcacaattgtaaacaccactacac
tgaacagcatcatccctatatccacatgtacttgttaacagaatacaattccctaggaagctggaatgctgga
agtcatggtgatgttctcatggttacagagaatctctctaaaaactaaaacctctttctgttttaccgcagTA
TGGAAGAATGTGAAGCTCTTTGCACTAGGATGGCAATCATGGTCAATGGAAGGTTTCAGGTGCCTTGGCAGTG
TCCAGCATCTAAAAATAGgtaataaagataatttcttgggtagtgcttagtgagaaggcttgatattta
ttcttttgtgagtatataaatgggtgcctctaaaaataaagggaataaaaactgagcaaaacagtatagtgga
agaattgagggctttgaagtcggaactgcattcaaatctgtctttaccatttactggttctgtgactcttgg
gcaagttacttaactactgtaagagttagtttccctggagatctacctcctagcttttgtgctatagatgaa
atgaaaaaaatttacatgtgccagtagtggtgagagcgcaagctttggagtcaaacacaaatgggtttgcat
cctggccctaccaattatgagctctgagccatgggcaagtgaactcctgggctcagtttctctgtaa
catctgtcagacttcatgggtccaggtgaggattaaaggagatcatgtatttacagcacatggcatggtgct
tcacataaaaataagttatttagtaaatgataactggttcttctctcagaaacttatttctgggctgccagg
ggccgcccctttttcatggcacaaagtgggttcccagggttcagttattcttttaaatagttttctggagatcc
tccatttgggtattttttcctgctttcagGTTTGGAGATGGTTATACAATAGTTGTACGAATAGCAGGGTCC
AACCCGGACCTGAAGCCTGTCCAGGATTTCTTTGGACTTGCATTTCTGGAAGTGTCTAAAAGAGAAACAC
CGGAACATGCTACAATACCAGCTTCCATCTTCATTATCTTCTCTGGCCAGGATATTTCAGCATCCTCTCCAG
AGCAAAAAGCGACTCCACATAGAAGACTACTCTGTTTCTCAGACAACACTTGACCAAgttaagctttgagtg
caaaacagatttacttctcaggggtgtggattcctgccccgacactcccgcccataggtccaagagcagtttg
tatcttgaattgggtgcttgaattcctgatctactattcctagctatgctttttactaaacctctctgaacct
gaaaagggagatgatgctatgtactctataggattattgtgagaatttactgtaataataaccataaaaaac
taccatttagtgagcacctaccatgggcccaggcattttacttgggtgcctaactcctatttaaattagataaaa
aagtaaccaaataggctcctgacacttaagaagtactcagtaaatattttcttccctcttcccttaatacaaga
ccgtatgtgcccagaagtaaatggatgactgagcagtggtgatgtaggggtggggggcgatatagaagtcag
tttttggccgggcgtgggtggctcatgcctgtaatcccagcactttgggaggctgaggagcagggcagatcatg

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

aggtcaggagatccagataatcctggccaacaggggtgaaaccccgctctctactaaaaatacaaaaattagct
gggcatgggtgggtgcgcactttagtcccagctacttgcgagggtgaggcaggagaattgctcgaacccagga
gggtggaggttacagtgagccaaggtctcgccactgcactccagcctggggacagagcaagacccccatttcaa
ggggggaaaaaaagtctatttttaagttggtattgcttttttcaagattcttccctccttcacacacagtt
ttctagttaatccatttatgtaattctgtatgctcctacttgacctaatttcaacatctggaaaaatagaac
tagaataaagaatgagcaagttgagtggtatttataaaggtccatcttaatcttttaacagGTATTTGTGAA
CTTTGCCAAGGACCAAAGTGATGATGACCCTTAAAAGACCTCTCATTACACAAAAACCAGACAGTAGTGGA
CGTTGCAGTTCTCACATCTTTTCTACAGGATGAGAAAGTGAAGAAAGCTATGTATGAAGAACTCTGTTTCTAT
ACGGGGTGGCTGAAAGTAAAGAGGAAGTAACTGGATACTGTACTGATACTATTCAATGCAATGCAATTCAATGCAAT
GAAAACAAAATTCCATTACAGGGGCAGTGCCTTTGTAGCCTATGTCTTGTATGGCTCTCAAGTGAAAGACTT
GAATTTAGTTTTTTTACCTATACCTATGTGAACTCTATTATGGAACCAATGGACATATGGGTTTGAAGTCA
CACTTTTTTTTTTTTTTTTTTGTTCCTGTGTATTCTCATTGGGGTTGCAACAATAATTCAAGTAATCATGG
CCAGCGATTATTGATCAAAATCAAAGGTAATGCACATCCTCATTACTAAGCCATGCCATGCCAGGAGAC
TGGTTTTCCCGGTGACACATCCATTGCTGGCAATGAGTGTGCCAGAGTTATTAGTGCCAAGTTTTTTCAGAAAG
TTTGAAGCACCATGGTGTGTCTGCTCACTTTTGTGAAAGCTGCTCTGCTCAGAGTCTATCAACATTGAATA
TCAGTTGACAGAATGGTGCCATGCGTGGCTAACATCCTGCTTTGATTCCCTCTGATAAGCTGTTCTGGTGGC
AGTAACATGCAACAAAAATGTGGGTGTCTCCAGGCACGGGAAACTTGGTTCCATTGTTATATTGTCTCTATGC
TTTCAGCCATGGGTCTACAGGGTCTATCCTTATGAGACTCTTAAATATACTTAGATCCTGGTAAGAGGCAAG
AATCAACAGCCAAACTGCTGGGGCTGCAAGCTGCTGAAGCCAGGGCATGGGATTAAAGAGATTGTGCGTTCA
AACCTAGGGAAGCCTGTGCCCATTGTCTGCTGACTGTCTGCTAACATGGTACACTGCATCTCAAGATGTTTAT
CTGACACAAGTGTATTATTTCTGGCTTTTTGAATTAATCTAGAAAATGAAAAGATGGAGTTGTATTTTGACA
AAAATGTTTGTACTTTTTAATGTTATTTTGGAATTTTAAAGTTCTATCAGTGACTTCTGAATCCTTAGAATGGC
CTCTTTGTGATACCCCTGTGGTATAGAGGAGTATGGCCACTGCCCACTATTTTTTATTCTTATGTAAGTTT
GCATCTCAGTCAGTACTAGTGCCTAGAAAAGCAATGTGATGGTCAAGATCTCATGACAZTTATATTGAGTTT
CTTTGAGATCATTAGGATACTCTTAATCTCACTTCATCAATCAAATATTTTTTGTAGTGTATGCTGTAGCTG
AAAGAGTATGTACGTACGTATAAGACTAGAGAGATATTAAGTCTCAGTACACTTCTGTGCCATGTTATTCA
GCTCACTGGTTTACAAATATAGGTTGTCTTGTGGTTGTAGGAGCCCACTGTAACAATACTGGGCAGCCTTTT
TTTTTTTTTTTTTAAATTGCAACAATGCAAAAGCCAAGAAAGTATAAGGGTCACAAGTCTAAACAATGAATTC
zTTCAACAGGGAAAAACAGCTAGCTTGAAAACCTTGCTGAAAAACACAACCTTGTGTTTATGGCATTTAGTACCT
TCAATAATTGGCTTTGAGATATTGGATACCCATTAAATCTGACAGTCTCAAATTTTTTCTCTCTTCAAT
CACTAGTCAAGAAAAATATAAAAAACAATACTTCCATATGGAGCATTTTTTCAGAGTTTTCTAACCAGT
CTTATTTTTCTAGTCAGTAAACATTTGTAAAAATACTGTTTCACTAATACTTACTGTTAACTGTCTTGAGAG
AAAAGAAAAATATGAGAGAACTATTGTTTGGGGAAGTTCAAGTGATCTTTCAATATCATTACTAACTTCTTC
CACTTTTTCCAGAATTTGAATATTAACGCTAAAGGTGTAAGACTTCAGATTTCAAATTAATCTTTCTATATT
TTTTAAATTTACAGAATATTATATAACCCACTGCTGAAAAAGAAAAAATGATTGTTTTAGAAGTTAAAGTC
AATATTGATTTTAAATATAAGTAATGAAGGCATATTTCCAATAACTAGTGATATGGCATCGTTGCATTTTAC
AGTATCTTCAAAAATACAGAATTTATAGAATAATTTCTCCTCATTTAATATTTTTCAAATCAAAGTTATGG
TTTCTCATTTTTACTAAAATCGTATTCTAATTTCTTATTATAGTAAATCTATGAGCAACTCTTCACTTTCGGT
TCTCTGATTTCAAGGCCATATTTTAAAAAATCAAAGGCAGTGTGAACATTTTTGAAGAAAAACAACATT
TTAATACAGATTGAAAGGACCTCTTCTGAAGCTAGAAAACAATCTATAGTTATACATCTTCATTAATACTGTG
TTACCTTTTTAAATAGTAATTTTTTACATTTTCTGTGTAAACCTAATTGTGGTAGAAATTTTTTACCAACTC
TATACTCAATCAAGCAAAATTTCTGTATATTCCCTGTGGAATGTACCTATGTGAGTTTCAAGAAATTTCTCAA
ATACGTGTTCAAAAATTTCTGCTTTTGCATCTTTGGGACACCTCAGAAAACCTATTAAACAACGTGAATATG
AGAAATACAGAAGAAAAATAAAGCCCTCTATACATAAATGCCAGCACAAATCATTGTTAAAAACAACCA
AACCTCACACTACTGTATTTTATTATCTGTACTGAAAGCAAATGCTTTGTGACTATTAAATGTTGCACATCA
TTCATTCACTGTATAGTAATCATTGACTAAAGCCATTTGTCTGTGTTTTCTTCTGTGGTTGTATATATCAG
GTAAAATATTTTCAAAGAGCCATGTGTCTGTAATACTGAACCACTTTGATATTGAGACATTAATTTGTAC
CCTTGTTATTATCTACTAGTAATAATGTAATACTGTAGAAATATTGCTCTAATTTCTTTTCAAATTTGTTGCA
TCCCCCTTAGAATGTTTCTATTTCCATAAGGATTTAGGTATGCTATTATCCCTTCTTATACCCTAAGATGAA
GCTGTTTGTGTCTCTTTGTTTCATCATTGGCCCTCATTCCAAGCACTTTACGCTGTCTGTAATGSGATCTAT
TTTTGCACTGGAATATCTGAGAAATTGMMAACTAGACAAAAGTTTCAACAAGATTTCTAAGTTAAATCAT
TTTCATWAAARRRRRAAARAAAAAAATTTTGTATGTCAATAACTTTATATGAAGTATTAAAAATGCATATT
TCTATGTTGTAATATAATGAGTCACAAAATAAAGCTGTGACAGTTCTGTTGGTCTACAGAAAttctatttg
tgcatttgtggccaccctactgttgaaggggtataaagccattagaaaagtagaggggaagtgtattggat
caaaaggaaaaatttagaaaagattcaaatgttcccttaatacataaaagagaactgaggggactacttgaa
aataaaaggttgttttgtattttcatgttgggttaagatactgagztaacwgggattaagtgttagaggtttt

tagataaatattctgcttaattgattatgaagctgcactgagatttctgaaaatgctctgtagctgagcttat
ttaataaatgttcacttgggtataggggaagctacaaaggcagccttcagtgtccttttgtttattcaaccaa
aaatataaggacacaaatgtagcagttatactgggaaggtgctgggggtgggtggcaatggtagcaggaaggc
gaagtagatatggaaacagaaatgataactaatatcggtgattccttccttttttctgtatagtgctgtg
cagacaacatatgagcagtgctgataaatgtaaatgtattkttcatagctcattaagaatcagtttcagaaa
gagatgtctgcttattttkgctrcttgaagaatccctgtcaaacagtccttttsaggaagtacaagaggctgt
ctctattttgtgacctcaggaatggctgtgacagtgctgtgagcagtccttttctgtggcacagatctgaac
tttgtgtgcagaaaaatccttggcttcaagtgaagcaagatgccccctgagcatcagcatcacaacttcatcc
tctatcttgaagttcatgttatagtgaactttaatgaaatcatagaacactgtttcttcgtgzaaacattgac
gagggagaggaaaaaactttattgaaaaataaaaaggcaggtaattagatgaaaaatgtttaccatgagg
ttttgtttttgtcttttgtttttgtttttgagaacacagaatctcgctctgtcgtccaggctggagtgcagcg
gcatgatcttggctcactgcaacctccgcctcccggttcaagcagattctcctcagcttcccaagtagctgg
tactacaggcatgcgccaccacacacagctaattttttgtatttttagtagagatgggggttctactatagctt
ggccaggctgggtctcaaactcctgacctaaagtgatccttctgccttgggctcccaaagtgctgggattaca
ggcatgagccaccttgcctggcctacctatgagccttgactaaaacattcttctatctgtagaaaagccca
aaagaacttttccagattcaaaaaacttggcactttgtaatggtaattgtttacattaaagtaaaaaa
aaaacccacttagcttcagttttcaagtgtttactgtgtgtcatgcacttcatttaattctcaacacctgc
cctatgaggtaaaaagtagctttttacatatgagtaaaattacagctcagtgagataagaaactcgtccaaagg
tacaggttcagtcagtgagggaggggttctttttgttgaagttaggtatcagttaaaattgaccttgtaaaa
tcacatcagcatcaatatataatatttaacaaatattttattgaactttactgtatgccagatacttctcta
ggtagtaggggttacaatgtagaagaaaatagaattcctgctctaattgaatttatatttcagtggtgaaaga
tgatgtgtggacaaacacatctaattgtattttgacagcaatgggtgctaagaagaaaataagacatggta
tgagataacaatggagagaagtcagagatggccttttctgaggagtgcagatgctcagaatcgaataaaaagag
caggagcagcctttcaaggtgggtgcgaaggatattcccgaagaaagaataagtggtgcgaagccccagtc
gggaaaaagcttggcaagctgaggtggtaaaagagctatctgtactaatgtctcttctgtactgagttagcaag
gacaagagtggcaggggggtggaactggagagatagcagggaccacatctcacaggatctcgccagcctttt
aaggtatttggatttttattgttagtgcaacaaggagcactggagagttttaaacagtagtggtgtgacctta
ctgtttcagaaagaacactggctactatgggaagaaaggacagtaggaagactaatagataatgggtggattg
aactaagatggtagcaacagatagggagctatggtaattgttcagtatccactttggagatagatccagcagg
acttgctgacagactgcagtgtaagggctgagggaaagagagctagcaaggtgacttctagtttgtgacctga
actaggtagatgggtgtgttaaaaatgcagtggttatttcagctgaagctcaggggcaatgatggaaaata
agataaatgtgtgcaagttggctgggctgggtgggtgcacgcctataatcacagcactttgggagggccgagac
gggtggatcacttgaggtcaggcgtttgcgaccagcctaacatgggtgaaaccccatctctaataaatacaaaa
aaaattagctgggcatgggtgggtgcagctgttaattccaagctacttgggagggctgagacaagagaatcgctt
gcacctgggagggcggaggtttcagtgagctgagatcgccaccattgcactccagcctgggcaacaagagcgaa
actccatctaaaaaaaagtggtgcaagttattgggtacgtatatatttaagcctcgggactgaatgagatta
cctagagagaatgaagataggggaacattcttgagcctgagatcctataacatttagaagagaagctagccag
ggaaattgagagggactggccagtgagagagagtaaaatctgaagagtataatgccatattgttgaagctcag
agaattattgtttcaggaaggcagaagttgtcaattatgtccaacattgctgagaagatgagtaaggaatgt
aatgagtaaaaggcaggaatggctattgaatttgttaagacagaggtccttgggtgaccttccctaaaggtcat
ttcagcatgacaaggatgcagacctaaactggaatagattttcaggggaagactgagaggaaggaaggggaga
ctgcaaatatatgcaaatctatatattttgtcttgaagggaaacagaaaaacaggacaacagctgggttacaga
acttttatttggattcttctccttgtctgttaataatataaggctgcacctgggcatcatcgtgcccttctctt
ttgattctgctgttttgccttttaactcctacaggcatttcttctggcaggagccttgtacacactctgggtgg
ttcatgaccaccaggacttgcctttcaaggcagaagcctgtgattcaaaaacttccattccttccaatcg
actattataaaactcattttgtctagtgtccatactcctttaccttaggatcctgggccaactgtcactaaa
gaaaacaaaccatttctcctcctgggggaatgtgcttctttcaagaaggtgcttagagaatagacattcta
ggtagacagcagttatgaaagaggtgtctctaattgaacctaacacttcttccaacatactccaggttcacctt
cctttgaagacagaaagtactacaacttcagttgccaccctagtcaggttcataggatttactcagtaggtt
aagagatgggtgctgagaatgaagtccagagtttgagtctacaagtggtcatttgcttctcctttgggacttct
gtggcaggggtgtggaaccaaactcagattgggtgcatactctgtttctccttctgaaaatacaattcaaatacc
ttcatagctctgtgacagcaacttataaccaacagacaatggattattttttattttattgatggcactcccc
atagccagaacttctactgagtcattcatcgtgtattccaggaatactttccaaatggagtagtgaggcact
tccaataacagttatgtaattctgttaacagtggtcctcaaatgacatgtgaaggaatgatgtacttcttttt
cagactcacagatctcatatgttaattcagtcacaagtagttagggcttatcacatgccaggcacaggttaa
tatgacaataatgaaaacaaaaccccttgcgtgtcatagaattcggtttctagtgggaagagacatacatata
caaatcatatataagtaaaacattaattatgggtgataacagaaaaagaaagcaggggagaggatagtgtct

tagttcaggttgctgtaacaagagtaccatagacaggggtggccttamacaacaaacatttctttctcacagat
ctagtggctgggaagtctaagaccaggggtgccagcatgaagaagttctgggtgaaggcttacattctgggttc
ctcacagtgagggctcttcacgcccttataaatgcactaattctattaccacatgacctcatctaaaacta
actcacaaggccaaacctcctatgttgagattaaggcttcaacataggaattttggggggacaaaaacatt
cagttctctagcagatagggggccactgtggggaggggattacaatcttaaatagggttgctcaggggaaggcttca
tgagggcacaagacctgaggagcatggaaatgatccacgcagatgcatgtctgggggaagagcattccaggga
gaggcagtgaggaaatggaccctcaggttaagagcatgttttagcaagttcagagaaaagcctcttaggtcagggtc
tgagagcgagagtggaataatgtaacaggacagatcatgtaaaagacttgtagaccacagtttagaggcttgga
ataactctgaggatgacatgaaggatattgggaggggggtgttgagcagaggagtgatgatctgggtttat
tttagattattctgactgctatgttgaaaaataggggttgagagcagaggggaagggcacaccagtttaaga
ccacttcccaaaaaatccagatgatgatgggcttgaaccaggagagttgctatagggatgtgaagcaatcaaa
ttctagatatttgggacctgctgacaggtacataaatgggtgtatgagagaaaaggcagtggtgagaatgactct
gaagttctgtggcctgagcaaaacagaaacctggagttgccaatagctgagataaagtttctgggttaaagatc
aagagctcagtttgggacaagaggggtgtgagatttctaataagacattcaggtggaggtgtggggaacacac
ctggatgtgatgagcctggaattcaaatccacatgtgtgaatttgaaaaaagattcagaaaagaggtgtgga
ctggaggagctatccacacacagatagcatttgaagccacaagactgggggtgatgacgcagggagtaagta
taaatagaaaagagaagtggtccaaggactcagtccttacatcagaggtcaaagagacaagatggagactag
gagatatgaaggcttatgcagtgagtggttcggagaaaacgaagaaacttccagaagaaaggagaggtcaac
tgctaacaagtcaagtgaatgagaactgggaaatgagttgacaagagcaatttggctggcgtgatcagaga
atgtttaagagataatgtgaggggaagaaaatggaggcaggaagtatgaactactcttttgagtcatttgta
aaggaaaaaaatgggggtgattttactcttttttttttaagataggaaaaaaagtatatatttatatgctg
aggagaagaatccagttgaaggggaataactgtggcttcaggagacaggaaaaactgcctaagccacatccttc
agcaagtgagtggtcaagtgagggtttgtctttgcttggggcacgggcggttcacccacagtgaaggaga
gaaggcagggcatgtgggttggtgttagaaagcgggcagacgggggtgggaacttatgaagactcttctgat
tggtattttctctgttaagtaaaaggaaaggtcattggctaagaatgaagatggacgcgcaggtgctgggag
tcatcgaagagaatgaaagaatgaataaaaatagagaaaaatacaatgctattatataaacctatacatttaa
attactttcctcttcaaattagacccccctggaaggcaaatctctgtcatgttaagtttaggaaaagtcata
atctacttggagctaagtatgaagaataaggtatatgatttaaacacataattctatttttgcttaaacg
gggaagcctttaaggtaacgtagagtgtaacctataaaaaagataaaaaatgtcagcttcccccttcccttaa
atatagtaaggctcaaaggtaattgatgttcataaaccttaagggtactataacagtcgaagcaaaatattaa
atgtcttctgtagtattttagttaacttttaattttattggtaatttatgtgtagttaaacacataacttggtg
cctcaactatttcacataaatgtgagatgtcttaggcctctttgaactctgcatggaaaagaatgctgagga
gggtatgtctgtttaagctgccactgctgctctcataaataatatctcacattgcataatagcccagataa
actgcagatcgcttagagcctgataaatgagaggaaactgtacagtttggttttccaaattggccttggacag
ctaattcacaatttactattttatagctgtaagattttaaaaatataaaaataacatagctgtaaatacagat
tatcagaagatagttctgtaaaaatgtaaatataatataagctgtaaatatacactatcggaagtaccaaactc
agtttttagtaggttagagggcatcaagtaaaatggaaatgaaagatgtttcaaggaaaactccaggacatctgt
ggcagttactaaagaaaccttcccttctcaggttcccagcatgctattttattgatgtaacaacattttcaata
aagttggtaaatatcactatattactgtcttatagcaacatagcaagagcttttagagatcatataagtatata
aaatgtgaattttaaaaaacaatgaatatgcaggatttttattagggaagcgtttccataaccataaatat
ttctttaaaacaaataaatgtcccaagatctctgttagtgatccaaactaagtagaaatttagtaaaattaat
tataaatgaacaatttcagcatataaaaccaacaagcttttctagatttttaacactgtgaccaatttgcat
tattttccaagttagaatgactaataatcaatgaatgtaaaagcaataattaatacagatgacattgtactt
ttccacagtaaaagaaataaacaatctaataatttttataaatccattttatatacacaataaacctttacta
agcaaatttttttaaatctcagggaactatagacatgatgaaaagatagatattttatataaataaattcaaa
aatactgtcagggaaggaaatgtaaaatccttatttgagttaaagaaaatgctataaagcaatgagttatca
aaatacagaagaggtatttctaaaacaaatgaaaaaccaagatgatgaaatagtgaactacttctaagtgtg
taacagatactgaaatgccagggtgaaagtgaactgaattattttcttaaagcagtgagagaatatgtaacttt
caaaaatgcaagaagcacagcaaatctaactaacttaccacctccttcaataaaaagcgagaacctcct
gggagaatttaagcaccattagcagacacatcttagagc

Figure 2A SEQ ID NO: 5

MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKAMPSAGTLPWVQ
GIICNANNPCFRYPTPGEAPGVVGNFNKSIVARLFSDDARRLLYSQKDTSMKDMRKVLRRTLQQIKKSSSNL
KLQDFLVDNETFSGFLYHNLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSE
LCGLPREKLAAAERVLRSNMDILKPIRLTLNSTSPFPKELAEATKLLHSLGTLAQELFSMRSWSDMRQE
VMFLTNNVSSSSSTQIYQAVSRIVCGHPEGGLKIKSLNWIYEDNNYKALFGGNGTEEDAETFYDNSTTPYC
NDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPATROVMAEVNKTQELAVFHDLEGMWHEELSPKIWTF
MENSQEMDLVRMLLDSRDNDHFWEQQLDGLDWTADQDIVAFLAKHPEDVQSSNGSVYTWREAFNETNQAIRT
ISRFMECVNLNKLEPIATEVWLINKSMELLDERKFWAGIVFTGITPGSIELPHHVKYKIRMDIDNVERTNK
IKDGYWDPGPRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGTEKKTGVYMQMPYPCYVDDIFLRVMSRS
MPLFMTLAWIYSVAVIIKGIYVEKEARLKETMRIMGLDNSILWFSWFISSLIPLLVSAGLLVILKLGNNL
PYSDPSVVFVFLSVFAVVTILQCFLISTLFSRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIFASL
LSPVAFGFGCEYFALFEEQIGVQWDNLFESPVEEDGFNLTSVSMMLFDTFLYGVMTWYIEAVFPGQYGI
PRPWYFPCTKSYWFGEESDEKSHPGSNQKRRISEICMEEEPHTLKLGVSIQNLVKVYRDGMKVAVDGLALNF
YEQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQNLGVCPQHNVLFDMLTVEEHI
WFIARLKLSEKHVKAEMEQLADVGLPSSKLKSKTSQLSGGMQRKLSVALAFVGGSKVILDEPTAGVDP
YSRRGIWELLLKYRQGRITILSTHMDVGLDRIAIISHGKLCVGSLLFLKNQLGTGYLLTLVKKDVE
SSLSSCRNSSSTVSYLKKEDSVSQSSSDAGLSDHESDTLTIDVSAISNLIRKHVSEARLVEDIGHELTYV
LPYEAKEGAFVELFHEIDRLSDLGISYGISETTLEEIFLKVAEESGVDAETSDGTLPARNRRAFGDK
QSLRPFTEDDAADPNDSIDPESRETDLLSGMDGKGSYQVKGWKLQQQFVALLWKRLLIARRSRKGFFA
QIVLPAVFVCIALVFLIVPPFGKYPSELQPMWYNEQYTFVSNDAPEDTGTLELLNALTDPGFGTRCME
GNPIPDTPCQAGEEWTAPVPQTIMDLFQNGNWTMNPSPACQCSSDKIKMLPVCPPGAGGLPPPQRKQ
NTADILQDLTGRNISDYLKTYVQIIAKSLKNKIWNVEFRYGGFSLGVSNTQALPPSQEVNDAIKQMKKHL
KLAJDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFLNVINNAILRANLQGENPSHYGITAFN
HPLNLTKQQLSEVALMTTSVDVLVSICVIFAMSFVPASFVFLIQERVSKAKHLQFISGVKPVYIWLNSFV
WDMCNYVVPATLVIIIFICFQKSYVSSTNLPLVALLLLLYGWSITPLMPASFVFKIPSTAYVVLTSVNL
FIGINGSVATFVLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSW
DLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQRILDGGGQNDILEIKEL
TKIYRRKRKPAVDRIICVGIIPGECFGLLVNGAGKSSTFKMLTGDTTVTRGDAFLNKNLSILSNIHEVHQN
GYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKVGWAIKRLGLVKYGEKYAGNYSNGNKRKLSTAMAL
IGGPPVFLDEPTTGMDPKARRFLWNCALSVVKEGRSVVLTSHSMEECEALCTRMAMVNGRFRCLGSVQH
LKNRFGDGYTIVVRIAGSNPDLKPVDQDFGLAFPGSVLKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRL
HIEDYSVSQTTLQVVFVNFQKQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV*

Figure 2B SEQ ID NO: 6

GTCCCTGCTGTGAGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGGGTG
CTGGCTGAGGGAACATGGCTTGTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCA
CTTTCAGAAGAAGACAAACATGTCAGCTGTTACTGGAAGTGGCCTGGCCTCTATTTATCT
TCCTGATCCTGATCTCTGTTTCGGCTGAGCTACCCACCCTATGAACAACATGAATGCCATT
TTCCAAATAAAGCCATGCCCTCTGCAGGAACACTTCCTTGGGTTCAGGGGATTATCTGTA
ATGCCAACAAACCCCTGTTTCCGTTACCCGACTCCTGGGGAGGCTCCCGAGTTGTTGGAA
ACTTTAACAAATCCATTGTGGCTCGCCTGTTCTCAGATGCTCGGAGGCTTCTTTTATACA
GCCAGAAAGACACCAGCATGAAGGACATGCGCAAAGTTCTGAGAACATTACAGCAGATCA
AGAAATCCAGCTCAAACCTGAAGCTTCAAGATTTCTGGTGGACAATGAAACCTTCTCTG
GGTTCCTGTATCACAACTCTCTCTCCCAAAGTCTACTGTGGACAAGATGCTGAGGGCTG
ATGTCATTCTCCACAAGGTATTTTTGCAAGGCTACCAGTTACATTTGACAAGTCTGTGCA
ATGGATCAAAATCAGAAGAGATGATTCAACTTGGTGACCAAGAAGTTTCTGAGCTTTGTG
GCCTACCAAGGGAGAACTGGCTGCAGCAGAGCGAGTACTTCGTTCCAACATGGACATCC
TGAAGCCAATCCTGAGAACACTAACTCTACATCTCCCTTCCCGAGCAAGGAGCTGGCTG
AAGCCACAAAAACATTGCTGCATAGTCTTGGGACTCTGGCCCAGGAGCTGTTTCAGCATGA
GAAGCTGGAGTGACATGCGACAGGAGGTGATGTTTCTGACCAATGTGAACAGCTCCAGCT
CCTCCACCCAAATCTACCAGGCTGTGTCTCGTATTGTCTGCGGGCATCCCGAGGGAGGGG
GGCTGAAGATCAAGTCTCTCAACTGGTATGAGGACAACAATAACAAAGCCCTCTTTGGAG
GCAATGGCACTGAGGAAGATGCTGAAACCTTCTATGACAACCTCTACAACCTCCTTACTGCA
ATGATTTGATGAAGAATTTGGAGTCTAGTCCTCTTTCCCGCATTATCTGGAAAGCTCTGA
AGCCGCTGCTCGTTGGGAAGATCCTGTATACACCTGACACTCCAGCCACAAGGCAGGTCA
TGGCTGAGGTGAACAAGACCTTCCAGGAACCTGGCTGTGTTCCATGATCTGGAAGGCATGT
GGGAGGAACTCAGCCCCAAGATCTGGACCTTCATGGAGAACAGCCAAGAAATGGACCTTG
TCCGGATGCTGTTGGACAGCAGGGACAATGACCACTTTTGGGAACAGCAGTTGGATGGCT
TAGATTGGACAGCCCAAGACATCGTGGCGTTTTTGGCCAAGCACCCAGAGGATGTCCAGT
CCAGTAATGGTTCTGTGTACACCTGGAGAGAAGCTTTCAACGAGACTAACCAGGCAATCC
GGACCATATCTCGCTTCATGGAGTGTGTCAACCTGAACAAGCTAGAACCCATAGCAACAG
AAGTCTGGCTCATCAACAAGTCCATGGAGCTGCTGGATGAGAGGAAGTTCTGGGCTGGTA
TTGTGTTCACTGGAATTACTCCAGGCAGCATTGAGCTGCCCCATCATGTCAAGTACAAGA
TCCGAATGGACATTGACAATGTGGAGAGGACAAATAAAATCAAGGATGGGTACTGGGACC
CTGGTCCTCGAGCTGACCCCTTTGAGGACATGCGGTACGTCTGGGGGGGCTTCGCCTACT
TGCAGGATGTGGTGGAGCAGGCAATCATCAGGGTCTGACGGGCACCGAGAAGAAAAC TG

GTGTCTATATGCAACAGATGCCCTATCCCTGTTACGTTGATGACATCTTTCTGCGGGTGA
TGAGCCGGTCAATGCCCCCTCTTCATGACGCTGGCCTGGATTTACTCAGTGGCTGTGATCA
TCAAGGGCATCGTGTATGAGAAGGAGGCACGGCTGAAAGAGACCATGCGGATCATGGGCC
TGGACAACAGCATCCTCTGGTTTAGCTGGTTCATTAGTAGCCTCATTCTCTTCTGTGA
GCGCTGGCCTGCTAGTGGTCATCCTGAAGTTAGGAAACCTGCTGCCCTACAGTGATCCCA
GCGTGGTGTGTTGTCTTCCTGTCCGTGTTTGCTGTGGTGACAATCCTGCAGTGCTTCCTGA
TTAGCACACTCTTCTCCAGAGCCAACCTGGCAGCAGCCTGTGGGGGCATCATCTACTTCA
CGCTGTACCTGCCCTACGTCTGTGTGTGGCATGGCAGGACTACGTGGGCTTCACACTCA
AGATCTTCGCTAGCCTGCTGTCTCCTGTGGCTTTTGGGTTTGGCTGTGAGTACTTTGCCC
TTTTTGAGGAGCAGGGCATTGGAGTGCAGTGGGACAACCTGTTTGAGAGTCCTGTGGAGG
AAGATGGCTTCAATCTCACCATTCTGGTCTCCATGATGCTGTTTGACACCTTCCTCTATG
GGGTGATGACCTGGTACATTGAGGCTGTCTTTCCAGGCCAGTACGGAATTCACAGGCCCT
GGTATTTTCTTGCACCAAGTCCTACTGGTTTGGCGAGGAAAGTGATGAGAAGAGCCACC
CTGGTTCCAACAGAGAATATCAGAAATCTGCATGGAGGAGGAACCCACCCACTTGA
AGCTGGGCGTGTCCATTTCAGAACCTGGTAAAAGTCTACCGAGATGGGATGAAGGTGGCTG
TCGATGGCCTGGCAGTGAATTTTTATGAGGGCCAGATCACCTCCTTCCTGGGCCACAATG
GAGCGGGGAAGACGACCACCATGTCAATCCTGACCGGGTTGTTCCCCCGACCTCGGGCA
CCGCCTACATCCTGGGAAAAGACATTCTGCTCTGAGATGAGCACCATCCGGCAGAACCTGG
GGGTCTGTCCCCAGCATAACGTGCTGTTTGACATGCTGACTGTGGAAGAACACATCTGGT
TCTATGCCCCGCTTGAAAGGGCTCTCTGAGAAGCACGTGAAGGCGGAGATGGAGCAGATGG
CCCTGGATGTTGGTTTGCCATCAAGCAAGCTGAAAAGCAAAACAAGCCAGCTGTCAGGTG
GAATGCAGAGAAAGCTATCTGTGGCCTTGGCCTTTGTGCGGGGATCTAAGGTGTCTATC
TGGATGAACCCACAGCTGGTGTGGACCCTTACTCCCGCAGGGGAATATGGGAGCTGCTGC
TGAAATACCGACAAGGCCGCACCATTATTCTCTCTACACACCACATGGATGAAGCGGACG
TCCTGGGGGACAGGATTGCCATCATCTCCCATGGGAAGCTGTGCTGTGTGGGCTCCTCCC
TGTTTCTGAAGAACCAGCTGGGAACAGGCTACTACCTGACCTTGGTCAAGAAAGATGTGG
AATCCTCCCTCAGTTCTGCAGAAACAGTAGTAGCACTGTGTCATACCTGAAAAAGGAGG
ACAGTGTTTCTCAGAGCAGTTCTGATGCTGGCCTGGGCAGCGACCATGAGAGTGACACGC
TGACCATCGATGTCTCTGCTATCTCCAACCTCATCAGGAAGCATGTGTCTGAAGCCCCGC
TGGTGGGAAGACATAGGGCATGAGCTGACCTATGTGCTGCCATATGAAGCTGCTAAGGAGG
GAGCCTTTGTGGAACCTTTTCATGAGATTGATGACCGGCTCTCAGACCTGGGCATTTCTA
GTTATGGCATCTCAGAGACGACCCTGGAAGAAATATTCCTCAAGGTGGCCGAAGAGAGTG

GGGTGGATGCTGAGACCTCAGATGGTACCTTGCCAGCAAGACGAAACAGGCGGGCCTTCG
GGGACAAGCAGAGCTGTCTTCGCCCCGTTCACTGAAGATGATGCTGCTGATCCAAATGATT
CTGACATAGACCCAGAATCCAGAGAGACAGACTTGCTCAGTGGGATGGATGGCAAAGGGT
CCTACCAGGTGAAAGGCTGGAACTTACACAGCAACAGTTTGTGGCCCTTTTGTGGAAGA
GACTGCTAATTGCCAGACGGAGTCGGAAAGGATTTTTTGTCTCAGATTGTCTTGCCAGCTG
TGTTTGTCTGCATTGCCCTTGTGTTTACAGCTGATCGTGCCACCCTTTGGCAAGTACCCCA
GCCTGGAACCTTACGCCCTGGATGTACAACGAACAGTACACATTTGTCTCAGCAATGATGCTC
CTGAGGACACGGGAACCTTGAACTCTTAAACGCCCTCACCAAAGACCCTGGCTTCGGGA
CCCGCTGTATGGAAGGAAACCAATCCCAGACACGCCCTGCCAGGCAGGGGAGGAAGAGT
GGACCACTGCCCCAGTTCCCCAGACCATCATGGACCTCTTCCAGAATGGGAACTGGACAA
TGCAGAACCCTTACCTGCATGCCAGTGTAGCAGCGACAAAATCAAGAAGATGCTGCCTG
TGTGTCCCCCAGGGGCAGGGGGCTGCCTCCTCCACAAAGAAAACAAAACACTGCAGATA
TCCTTCAGGACCTGACAGGAAGAAACATTTTCGGATTATCTGGTGAAGACGTATGTGCAGA
TCATAGCCAAAAGCTTAAAGAACAAGATCTGGGTGAATGAGTTTAGGTATGGCGGCTTTT
CCCTGGGTGTCAGTAATACTCAAGCACTTCTCCGAGTCAAGAAGTTAATGATGCCATCA
AACAAATGAAGAAACACCTAAAGCTGGCCAAGGACAGTTCTGCAGATCGATTTCTCAACA
GCTTGGGAAGATTTATGACAGGACTGGACACCAGAAATAATGTCAAGGTGTGGTTCAATA
ACAAGGGCTGGCATGCAATCAGCTCTTTCTGAATGTCATCAACAATGCCATTCTCCGGG
CCAACCTGCAAAAGGGAGAGAACCCTAGCCATTATGGAATTACTGCTTTCAATCATCCCC
TGAATCTCACCAAGCAGCAGCTCTCAGAGGTGGCTCTGATGACCACATCAGTGGATGTCC
TTGTGTCCATCTGTGTCATCTTTTGCAATGTCTTCGTCCCAGCCAGCTTTGTCTGATTCC
TGATCCAGGAGCGGGTCAGCAAAAGCAAAACACCTGCAGTTCATCAGTGGAGTGAAGCCTG
TCATCTACTGGCTCTCTAATTTTGTCTGGGATATGTGCAATTACGTTGTCCCTGCCACAC
TGGTCATTATCATCTTCATCTGCTTCCAGCAGAAGTCCATATGTGTCCTCCACCAATCTGC
CTGTGCTAGCCCTTCTACTTTTGTCTGTATGGGTGGTCAATCACACCTCTCATGTACCCAG
CCTCCTTTGTGTTCAAGATCCCCAGCACAGCCTATGTGGTGTCTACCAGCGTGAACCTCT
TCATTGGCATTAAATGGCAGCGTGGCCACCTTTGTGCTGGAGCTGTTACCGACAATAAGC
TGAATAATATCAATGATATCCTGAAGTCCGTGTTCTTGATCTTCCCACATTTTGCCTGG
GACGAGGGCTCATCGACATGGTGAAAAACCAGGCAATGGCTGATGCCCTGGAAAGGTTTG
GGGAGAATCGCTTTGTGTCACCATTATCTTGGGACTTGGTGGGACGAAACCTCTTCGCCA
TGGCCGTGGAAGGGGTGGTGTCTTTCCTCATTACTGTTCTGATCCAGTACAGATTCTTCA
TCAGGCCCAGACCTGTAAATGCAAAGCTATCTCCTCTGAATGATGAAGATGAAGATGTGA
GGCGGGAAAGACAGAGAATTCTTGATGGTGGAGGCCAGAATGACATCTTAGAAATCAAGG
AGTTGACGAAGATATATAGAAGGAAGCGGAAGCCTGCTGTTGACAGGATTTGCGTGGGCA

1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000 12000 13000 14000 15000 16000 17000 18000 19000 20000 21000 22000 23000 24000 25000 26000 27000 28000 29000 30000 31000 32000 33000 34000 35000 36000 37000 38000 39000 40000 41000 42000 43000 44000 45000 46000 47000 48000 49000 50000 51000 52000 53000 54000 55000 56000 57000 58000 59000 60000 61000 62000 63000 64000 65000 66000 67000 68000 69000 70000 71000 72000 73000 74000 75000 76000 77000 78000 79000 80000 81000 82000 83000 84000 85000 86000 87000 88000 89000 90000 91000 92000 93000 94000 95000 96000 97000 98000 99000 100000

TTCCTCCTGGTGAGTGCTTTGGGCTCCTGGGAGTTAATGGGGCTGGAAAATCATCAACTT
TCAAGATGTTAACAGGAGATACCACTGTTACCAGAGGAGATGCTTTCCTTAACAAAAATA
GTATCTTATCAAACATCCATGAAGTACATCAGAACATGGGCTACTGCCCTCAGTTTGATG
CCATCACAGAGCTGTTGACTGGGAGAGAACACGTGGAGTTCTTTGCCCTTTTGAGAGGAG
TCCCAGAGAAAGAAGTTGGCAAGGTTGGTGAGTGGGCGATTTCGGAACTGGGCCTCGTGA
AGTATGGAGAAAAATATGCTGGTAACTATAGTGGAGGCAACAAACGCAAGCTCTCTACAG
CCATGGCTTTTGATCGGCGGGCCTCCTGTGGTGTCTTCTGGATGAACCCACCACAGGCATGG
ATCCCAAAGCCCGGCGGTTCTTGTGGAATTGTGCCCTAAGTGTTGTCAAGGAGGGGAGAT
CAGTAGTGCTTACATCTCATAGTATGGAAGAATGTGAAGCTCTTTGCACTAGGATGGCAA
TCATGGTCAATGGAAGGTTTCAGGTGCCTTGGCAGTGTCCAGCATCTAAAAATAGGTTTG
GAGATGGTTATACAATAGTTGTACGAATAGCAGGGTCCAACCCGGACCTGAAGCCTGTCC
AGGATTTCTTTGGACTTGCATTTCTGGAAGTGTTCTAAAAGAGAAACACCGGAACATGC
TACAATACCAGCTTCCATCTTCATTATCTTCTCTGGCCAGGATATTTCAGCATCCTCTCCC
AGAGCAAAAAGCGACTCCACATAGAAGACTACTCTGTTTCTCAGACAACACTTGACCAAG
TATTTGTGAACTTTTGCCAAGGACCAAAGTGATGATGACCACTTAAAAGACCTCTCATTAC
ACAAAAACCAGACAGTAGTGGACGTTGCAGTTCTCACATCTTTTCTACAGGATGAGAAAG
TGAAAGAAAGCTATGTATGAAGAATCCTGTTTCATACGGGGTGGCTGAAAGTAAAGAGGAA
CTAGACTTTTCTTTGCACCATGTGAAGTGTTGTGGAGAAAAGAGCCAGAAGTTGATGTGG
GAAGAAGTAACTGGATACTGTACTGATACTATTCAATGCAATGCAATTCAATGCAATGA
AAACAAAATTCCATTACAGGGGCAGTGCCTTTGTAGCCTATGTCTTGTATGGCTCTCAAG
TGAAAGACTTGAATTTAGTTTTTTTACCTATACCTATGTGAACTCTATTATGGAACCCAA
TGGACATATGGGTTTGAACTCACACTTTTTTTTTTTTTTTTTTGTTCCTGTGTATTCTCATT
GGGGTTGCAACAATAATTCATCAAGTAATCATGGCCAGCGATTATTGATCAAAATCAAAA
GGTAATGCACATCCTCATTCACTAAGCCATGCCATGCCCAGGAGACTGGTTTTCCCGGTGA
CACATCCATTGCTGGCAATGAGTGTGCCAGAGTTATTAGTGCCAAGTTTTTCAGAAAGTT
TGAAGCACCATGGTGTGTATGCTCACTTTTGTGAAAGCTGCTCTGCTCAGAGTCTATCA
ACATTGAATATCAGTTGACAGAATGGTGCCATGCGTGGCTAACATCCTGCTTTGATTCCC
TCTGATAAGCTGTTCTGGTGGCAGTAACATGCAACAAAAATGTGGGTGTCTCCAGGCACG
GGAACTTGTTCCATTGTTATATTGTCCTATGCTTCGAGCCATGGGTCTACAGGGTCAT
CCTTATGAGACTCTTAAATATACTTAGATCCTGGTAAGAGGCAAAGAATCAACAGCCAAA
CTGCTGGGGCTGCAACTGCTGAAGCCAGGGCATGGGATTAAAGAGATTGTGCGTTCAAAC
CTAGGGAAGCCTGTGCCCATTGTCTGCTGCTAACATGGTACACTGCATCTCAA
GATGTTTATCTGACACAAGTGTATTATTCTGGCTTTTTTGAATTAATCTAGAAAATGAAA

Figure 3

Promoter, 8797 bp

Distances numbered using first base of promoter as -1

Name	Pos. of 1st base in sense strand	Hit Site	% Match	Strand	SEQ ID No.
LXRE		Target: AGGTCA (NNNN){AGGTCA			7
DR4	-7531	AGAGGCAGGTGGATCATTTGAGGTCA	88	sense	8
DR4	-5085	TTGAGCGGGTGATCACTTGAGGTCA	88	antisense	9
DR4	-4389	CAAGCGGGCAGATCACTTGAGGTTA	88	antisense	10
DR4	-1641	CAAGTGGGCAGCTCACCTCAGGTCA	94	antisense	11
DR1		None			
PPAR		Target: NNNNN(A)NN(T)TGACCT(N/NN)TGACCT			12
DR2	-7718	CTTTGA(A)GC(C)TGATCATATGACCT	88	antisense	13
DR2	-7521	AGGCTG(G)TC(T)CGAACTCCTGACCT	88	antisense	14
DR2	-5708	CTTAAT(T)GG(T)GGWGTGTTGACCT	91	antisense	15
DR2	-2894	CAGGAT(G)GC(G)TAAACTCCTGACCT	88	antisense	16
DR2	-1649	AGGTTG(G)TT(T)CGAACTCCTGACCT	88	sense	17
DR2	-1140	TCAAGG(T)AG(G)AGACCTTGTGGCCT	88	sense	18
DR1		None			

Name	Pos. of 1st base in sense strand	Hit Site	% Match	Strand
SREBP		Target: ATCACCCCAC		
	-8523	GAGATGTGCTATGACCCCAC	90	antisense
	-3651	GTGAGCCCAGATCACACCAC	90	antisense
	-7747	TCCATCCATCCACACCCCAC	80	antisense
	-5485	CCCTTTTATTACACCTCAC	80	antisense
	-5248	GTAAGCCCAAGATCATGCCAC	80	antisense
	-5073	ACCTCAAGTGATCACCCGCC	80	sense
	-2252	GGCTCAAGCGATCTTCCCAC	80	antisense
	-2209	CCATGATTGGATCATGTGCAC	80	sense
	-1794	GTGAGTCGAGATCATGCCAC	80	antisense
	-519	TGCTTTTGTTTTTCCGCCAC	80	antisense
	-478	CCGCCTTCCCTCACCCCCAG	80	sense
-158	ACCCTCCACCCCACCCCCAC	80	sense	

ROR	Target: $(w)\{0,8\}$ WRGGTCA	
-8435	CTGGGCAAGGATGGGTCA	100 sense
-8434	TGGGCAAGGATGGGTCA	100 sense
-7025	AAAAAGCACCAAGGTCA	100 antisense
-3989	AGAAAGATGCCAGGGTCA	100 sense
-2638	GAGGAGATGGAGGGTCA	100 sense
32		
33		
34		
35		
36		
37		

Exon 1, 303 bp

Distances numbered using start of Exon 1 as + 1

Name	Pos. of 1st base in +	Hit Site	% Match	Strand	Q ID NO.
LXRE		Target: AGGTCA (NNNN)AGGTCA			7
DR4	4	CCGAGCGCAGAGGTTACTATCGGTCA	92	antisense	38
DR1		None			

PPAR

DR2

DR1

SREBP

ROR

5' Intron 1, 930 bp

Positions numbered using the first position in intron 1 as + 1

Name	Pos. of 1st base in +	Hit Site	Matc	Strand	SEQ ID N
LXRE		Target: AGGTCA (NNNN)AGGTCA			7
DR4	458	GCCAATTCCCAAGTCAAGACAGACCA	88	antisense	39
DR1		None			
PPAR		Target: NNNNN (A) NN (T) TGACCT (N/NN) TGACCT			12
DR2		None			
DR1		None			
SREBP		Target: ATCACCCAC			19
	326	GGACCTGCAGTCTCCCCAC	80	antisense	40
ROR		Target: (W) {0,8}WRGGTCA			32
	17	AACGCCCAAGTAAGTCA	94	antisense	41
	161	GAGCTCGTACTAGGACA	94	antisense	42
	181	GCAGAGTCTGGGTCA	94	antisense	43
	181	CGAGAGTCTGGGTCA	94	antisense	44
	478	AGCCAATTCACAGGTCA	94	antisense	45
	559	ACGGACCGTTTGGGACA	94	antisense	46
	559	CACGGACCGTTTGGGACA	94	antisense	47
	559	CCACGACCGTTTGGGACA	94	antisense	48
	589	ACTAGAGGCTTGGGTCT	94	sense	49
	590	CTAGAGGCTTGGGTCT	94	sense	50
	612	CCCTACCCCTCAGGTCA	94	antisense	51
	612	TCCCTACCCCTCAGGTCA	94	antisense	52
	668	GSTCTGCCCGGAGACA	94	antisense	53
	864	TTTTAGTGAGANGGTTA	94	sense	54

3' end of Intron 1, 12504 bp

Positions numbered using the first base 15 to the start of Exon 2 as -1

Name	Pos. of 1st base in +	Hit Site	Mat	Strand
LXRE		Target: AGGTCA (NNNN)AGGTCA		7
DR4	-7188	TGAGCAGTAGATCACTTGAGGTCA	93	sense
DR4	-11050	CGAGCTGGCGGATCACCCTGAGGTCA	86	sense
DR4	-7670	AAGCTAACAAGGTTACTGAAGGCCA	86	antisense
DR4	-4696	AGAGGTGGCGGATCACCCTGAGGTCA	86	antisense
DR1		None		

PPAR		Target: NNNNN (A) NN (T) TGACCT (N/NN) TGACCT		12
DR1		None		
DR2	-10281	CTCGAT (T) TC (C) TGACCTCGTGATCC	86	antisense
DR2	-5996	CAAAAC (A) TT (G) TGCCCTTTTGAAC	86	antisense
DR2	-932	GCGCTA (G) GG (T) TGTCCTCATTACCT	86	sense
DR2	-597	CTCGAT (T) TC (T) TGACCTCGTGATCC	86	sense

SREBP		Target: ATCACCCAC		19
	-7009	GTGAGCTGAGATCACACCAC	90	sense
	-11869	TTCAAGGATGATCACACAT	80	antisense
	-11616	GGCTCAAGTGATCCTCCAC	80	antisense
	-10100	GTGAGCCGAGATCGGCCAC	80	sense
	-8584	GTGAGTTATGATCATGCCAC	80	antisense
	-5591	CCACTGTTTCAACAACCCAC	80	sense
	-4684	ACCTCAGGATATCGGCCAC	80	sense
	-4128	AAATGTGACAAATCTCCACAC	80	antisense
	-2524	AAATAGAAATATCACCTCCC	80	antisense
	-1577	CCTTTTATCTACCACCCAC	80	antisense

Coding Polymorphisms (cSNP):																			
Exon	Numbering based on Pullinger et al., 2000	Numbering based on AJ012376.1	Sequence	Change in codon	Coding Nucl. Change	AA change	Genotypes	ALA	RLA	TDW	TDS	NL16	NL27	QC11	QC12	CZ01	JP01	NL16-212	BC11
Exon 0/5'UT	69	NA	GGAAGTGTCCGGCAAAA GGAAGTGTCTCGGCAAAA	NA	C69T	Not applicable	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
Exon 0/5'UT	127	NA	CCGGACCCGACAGCC CCGGACCCGACAGCC	NA	C127G	Not applicable	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
Exon 1/5'UT	319	(-) 256	ACACGCCAGGCGTCCTGC ACACGCCAGGCGTCCTGC	NA	InsG319	Not applicable	wt	wt	wt	wt	het	wt	het	wt	wt	wt	wt	wt	wt
Exon 1/5'UT	378	(-) 198	ACAGCTGGGGTGTCTGGCTG ACAGCTGGGGTGTCTGGCTG	NA	G378C	Not applicable	wt	wt	wt	wt	het	wt	het	wt	wt	wt	wt	wt	wt
Exon 5	869	414	CTGGGTTCTGTATCACAACC CTGGGTTCTGTATCACAACC	CTG-CTA	G869A	No aa change	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
Exon 6	1051	596	GGCCTACCAAGGAGAAACTG GGCCTACCAAGGAGAAACTG	AGG-AAG	G1051A	R219K	het	het	het	G/G	het	het	het	het	het	G/G	G/G	het	het
Exon 8	1331	876	GCGGGCATCCGAGGGAGGGG GCGGGCATCTGAGGGAGGGG	CCC-CCT	C1331T	No aa change	het	wt	het	wt	wt	het	wt	wt	wt	wt	wt	het	wt
Exon 8	1343	888	AGGAGGGGGGCTGAAGATCA AGGAGGGGGACTGAAGATCA	GGG-GGA	G1343A	No aa change	wt	het	wt	wt	het	wt	wt	wt	het	wt	wt	wt	het
Exon 10	1591	1136	TGACTCCAGGTGAACAAGACC TGACTCCAGGCGAACAAGACC	GTG-GGG	T1591C	V399A	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 15	2706	2251	GCAGGACTACGTGGGCTTCAC GCAGGACTACATGGGCTTCAC	GTG-ATG	G2706A	V771M	wt	wt	wt	wt	wt	wt	wt	wt	het	wt	wt	wt	wt
Exon 15	2715	2260	CGTGGGCTTCACACTCAAGAT CGTGGGCTTCGACTCAAGAT	ACA-CCA	A2715C	T774P	wt	wt	wt	wt	wt	wt	wt	het	wt	wt	wt	wt	wt
Exon 15	2723	2268	TCACACTCAAGATCTTCGCTg TCACACTCAAGATCTTCGCTg	AAG-AAC	G2723C	K776N	wt	het	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 16	2868	2413	CCACTTCGGTCTCCATG CCACTTCGATCTCCATG	GTG-ATC	G2868A	V825I	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	het	wt	wt
Exon 17	3044	2589	GAAGAGAATAATCAAGAG GAAGAGAATGTCAAGAG	ATA-ATG	A3044G	I883M	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	het	wt	G
Exon 21	3554	3099	GATCTAAGGTGTTCATCTGG GATCTAAGGTGGTTCATCTGS	GTT-GTG	T3554G	No aa change	het	wt	wt	wt	wt	wt	wt	het	wt	wt	wt	wt	wt
Exon 23	3911	3456	GGCACCATTGAGGTGACACG GGCACCATTGACGTGACACG	GAG-GAC	G3911C	E1172D	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 34	5155	4700	CTGGACACCAAAATATGTC CTGGACACCAAAATATGTC	AGA-AAA	G5155A	R1587K	wt	het	het	wt	wt	het	wt	wt	het	wt	het	het	het
Exon 37	5587	5132	TCCTATGTGTCTCCACCAAT	TCC-TGC	C5587G	S1731C	het	wt	wt	wt	wt	wt	wt	het	wt	wt	wt	wt	wt

Non-coding Polymorphisms (SNP):																							
Intron	Numbering based on Pullinger et al., 2000	Numbering based on AJ012376.1	Sequence	Change in non-coding	Coding Nucl. Change	AA change	Genotypes	ABE	MGA	ALA	RLA	TOW	TDS	NL16	NL27	QC11	QC12	CZ01	JP01	NL16-212	NL20	BC11	BAC
Promotor	(-) 191	NA	CGGGGAAGGGACGACGACGG CGGGGAAGGGACGACGACGG	G to C	Not applicable	Not applicable	CG	CG	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G
Promotor	(-) 17	NA	GGCCGGGAAGGGGGCGG GGCCGGGAAGGGGGCGG	C to G	Not applicable	Not applicable	CG	CG	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G	C/G
Intron 0	(-) 1163	(-) 1435	AGTATCCCTGTGTTACAGAA AGTATCCCTCCCTGTTCAGGAA	CCCT Insertion	Not applicable	Not applicable						no ins		no ins									
Intron 0	(-) 1095	(-) 1367	GTGACACCCAGCGAGTAGGG GTGACACCCAGCGGAGTAGGG	A to G	Not applicable	Not applicable						A/A		A/A									
Intron 0	(-) 1027	(-) 1299	TATGTGCTGACCGTGGAGCTTGT TATGTGCTGACCGTGGAGCTTGT	G to A	Not applicable	Not applicable						A/A		A/A									
Intron 0	(-) 720	(-) 992	CCTCGGCTGCCAGGTTACGCAATT CCTCGGCTCGCGGGTTCAGCGAATT	G to A	Not applicable	Not applicable						A/A		A/A									
Intron 0	(-) 461	(-) 733	GAAATTAGTATGTAAGGAAG GAAATTAGTCTGTMAAGGAAG	A to C	Not applicable	Not applicable						A/A		A/A									
Intron 0	(-) 362	(-) 634	CATTTCTAGAAAAGAGAGGT CATTTCTAGGAAGAGAGGT	A to G	Not applicable	Not applicable						A/A		A/A									
Intron 7	(+) 2383	(+) 2383	TTTAAAGGGGKTATTAGGA	Frequencies unknown: G and T	Not applicable	Not applicable																	
Intron 7	(+) 3035	(+) 3035	GAAGAAATTTKTTTTTIGATT	Frequencies unknown: G and T	Not applicable	Not applicable																	
Intron 7	(-) 15	(-) 15	TCTGTCCCATCCCTGACG TCTGTCCCAATCCCTGACG	A insertion	Not applicable	Not applicable			wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Intron 9	(-) 42	(-) 42	AGGAGCCAAACGCTCATTTG AGGAGCCAAACGCGCTCATTTG	G insertion	Not applicable	Not applicable			het	het	het	wt	het	het	wt	wt	wt	wt	wt	wt	wt	wt	wt
Intron 13	(+) 24	(+) 24	AAGCCACTGTGTTTTAACCACT AAGCCACTGTATTTAACCACT	T to A	Not applicable	Not applicable			wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Intron 13	(-) 83	(-) 83	GCTCCCTCTAGCATGCAGGCTC GCTCCCTCTAGTATGAGGCTC	C to T	Not applicable	Not applicable																	
Intron 15	(-) 4-5	(-) 4-5	TTGCCGTGTTTCTACAGAGCC TTGCCGTGTTCTCAGAGCC	CA deletion	Not applicable	Not applicable																	
Intron 17	(+) 2000	(+) 2000	GCGCAGTGCSCGTGTGTCCTTA	Frequencies unknown: C and G	Not applicable	Not applicable																	
Intron 21	(+) 118	(+) 118	CTCTTCTGTTAKCAGAGAGAGA	Frequencies unknown: G and T	Not applicable	Not applicable																	

																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										</
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Figure 5A

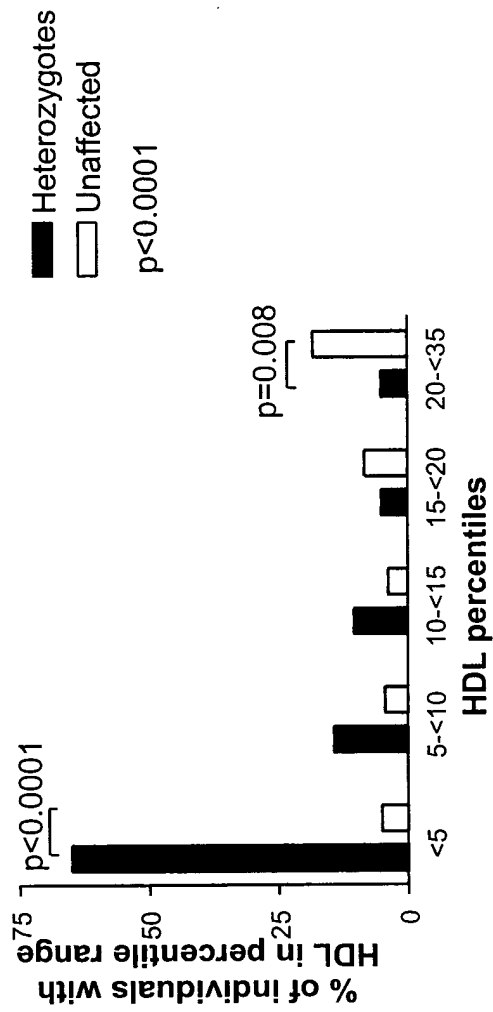


Figure 5B

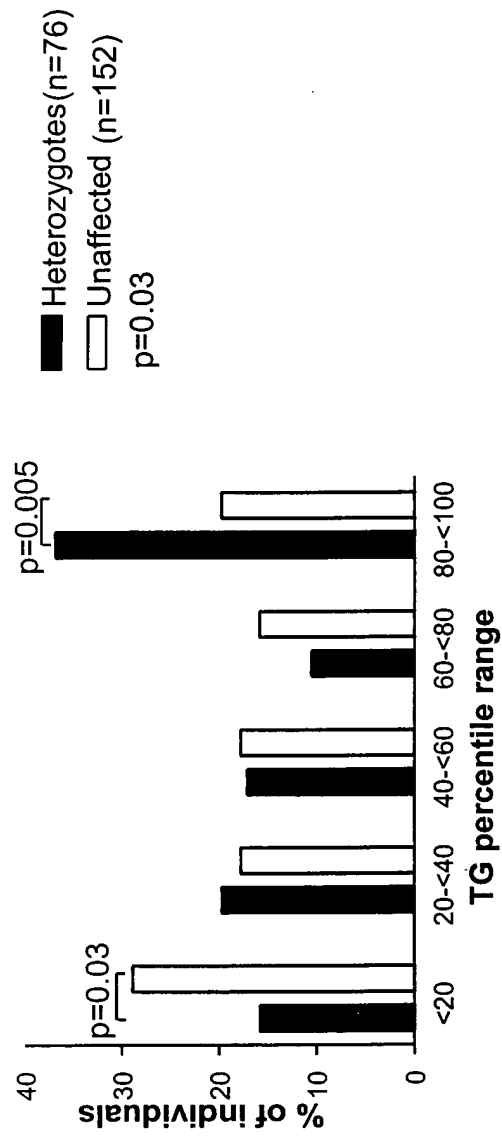


Figure 6

	TD Patients	Heterozygotes	Unaffected family members	P-value heterozygotes vs. unaffected	P-value TD patients vs. unaffected
number	5	77 ^A	156 ^A		
age (yrs) range	43.4±9.0 31-56	42.5±19.6 5-81	39.9±21.0 4-86	0.35	0.71
m/f	3/2	33/44	82/74	0.16	0.74
TC (mmol/L)	2.34±1.03	4.52±1.12	4.71±1.07	0.23	<0.0001
TG (mmol/L)	1.95±0.97	1.66±1.59	1.20±1.03	0.03	0.11
HDL (mmol/L)	0.08±0.05	0.74±0.24	1.31±0.35	<0.0001	<0.0001
LDL (mmol/L)	1.37±1.02	3.03±0.99	2.84±0.87	0.171	0.0003
ApoA-I (g/L)	0.03±0.04 (3)	0.92±0.32 (61)	1.43±0.26 (55)	<0.0001	<0.0001
ApoA-II (g/L)	0.10±0.08 (2)	0.35±0.08 (46)	0.39±0.08 (43)	0.01	<0.0001
ApoB (g/L)	0.89±0.53 (2)	0.93±0.25 (52)	0.94±0.33 (42)	0.88	0.84
CHD ≥ 20 yrs	20% (1/5)	12.9% (8/62)	4.1% (5/122)	0.03	0.10
Odds Ratio (95% CI)				3.47 (1.08-11.09)	5.85 (0.55-62.4)
Age of onset	38	48.9±8.6	60.4±12.8	0.08	

^A For TC, TG, LDL n=76 for heterozygotes, 153 for unaffected family members

Figure 7

Individual	Mutation	exon	disease (age of onset)
TD proband			
TD1	C1477R, ivs24+1G-->C	30, intron 24	CHD (38)
ABC1 heterozygotes			
TD4-201	unidentified	-	MI (<58)
FHA5-215	M1091T	22	MI (61)
FHA5-303	M1091T	22	CHD (<45)
TD1-363	C1477R	30	MI (51)
FHA3-301	Del(E,D) 1893,94	41	PVD (<54)
FHA3-305	Del(E,D) 1893,94	41	CHD (44)
FHA6-201	P2150L	48	CVA (36), fatal MI (58)
FHA2-301	R2144X	48	CAD (42), PTCA (47), femoral angioplasty (48), CABG (<50)
Unaffected family members			
FHA5-212	none	-	AP (62)
TD3-109	none	-	TIA (80)
FHA2-315	none	-	MI (51)
TD1-205	none	-	MI (62)
TD1-216	none	-	AP (47)

Figure 8

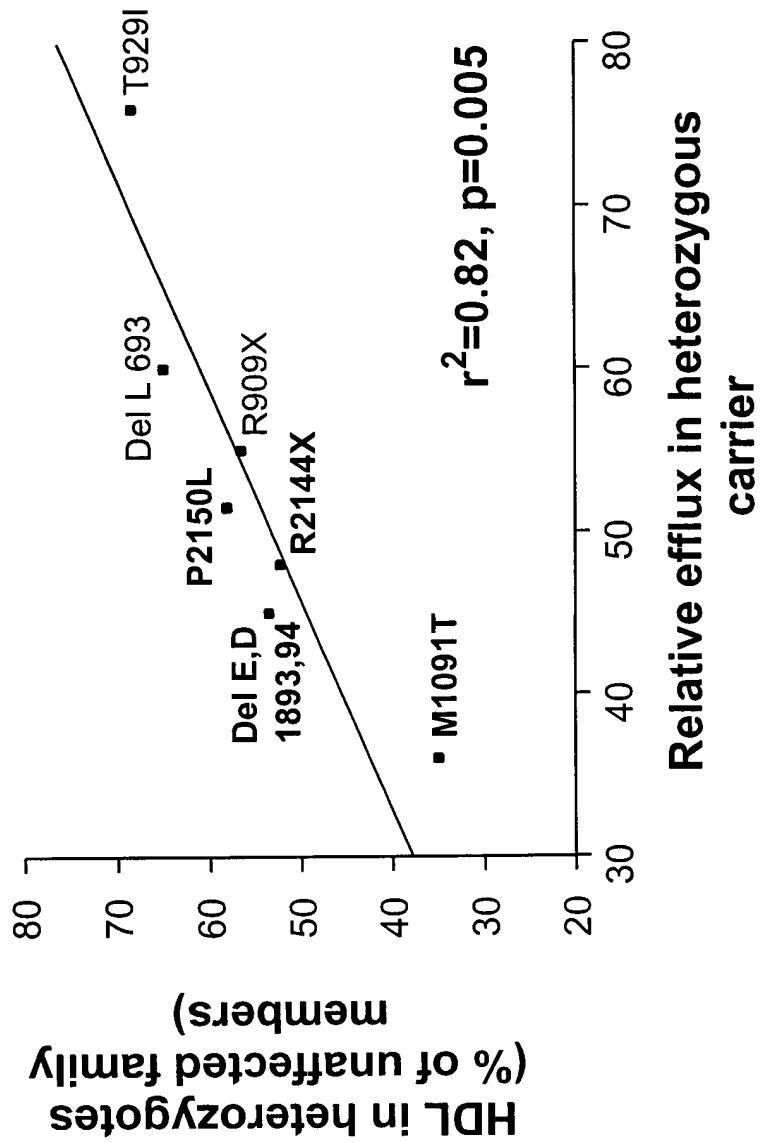


Figure 9

Figure 9

family	mutation	HDL in		HDL in unaffected		HDL in	Age and sex matched		CAD in	
		heterozygotes	mean + SD (n)	family members	heterozygotes		heterozygotes	population median ^A		heterozygotes
FHA1	Del L 693	0.79±0.20 (8)		1.22±0.35 (11)		64.8		1.39±0.08	-	
FHA2	R2144X	0.56±0.23 (12)		1.07±0.22 (20)		52.3		1.34±0.19	+	
FHA3	Del E,D 1893,94	0.77±0.24 (8)		1.44±0.38 (9)		53.5		1.30±0.17	+	
FHA4	R909X	0.59±0.26 (5)		1.04±0.37 (9)		56.5		1.39±0.24	-	
FHA5	M1091T	0.48±0.48 (4)		1.37±0.43 (6)		35.0		1.56±0.05	+	
FHA6	P2150L	0.61±0.07 (7)		1.05 (1)		58.1		1.30±0.22	+	
TD1	ivs25+1G-->C	0.78±0.06 (4)		1.35±0.29 (70)		57.8		1.22±0.22	-	
TD4	del C 6825-->2145X	0.91±0.10 (2)		1.00±0.05 (3)		91.0		1.31±0.16	-	
TD5	CTC6952-4TT-->2203X	0.80±0.20 (3)		1.65 (1)		48.5		1.39±0.19	-	
TD1	C1477R	0.82±0.18 (9)		1.35±0.29 (70)		60.7		1.37±0.14	+	
TD2	Q597R	0.82±0.07 (5)		none available		-		1.39±0.17	-	
TD3	T929I	1.01±0.18 (8)		1.48±0.42 (26)		68.2		1.33±0.19	-	
TD4	unidentified	0.74±0.05 (2)		1.00±0.05 (3)		73.5		1.49±0.09	+	

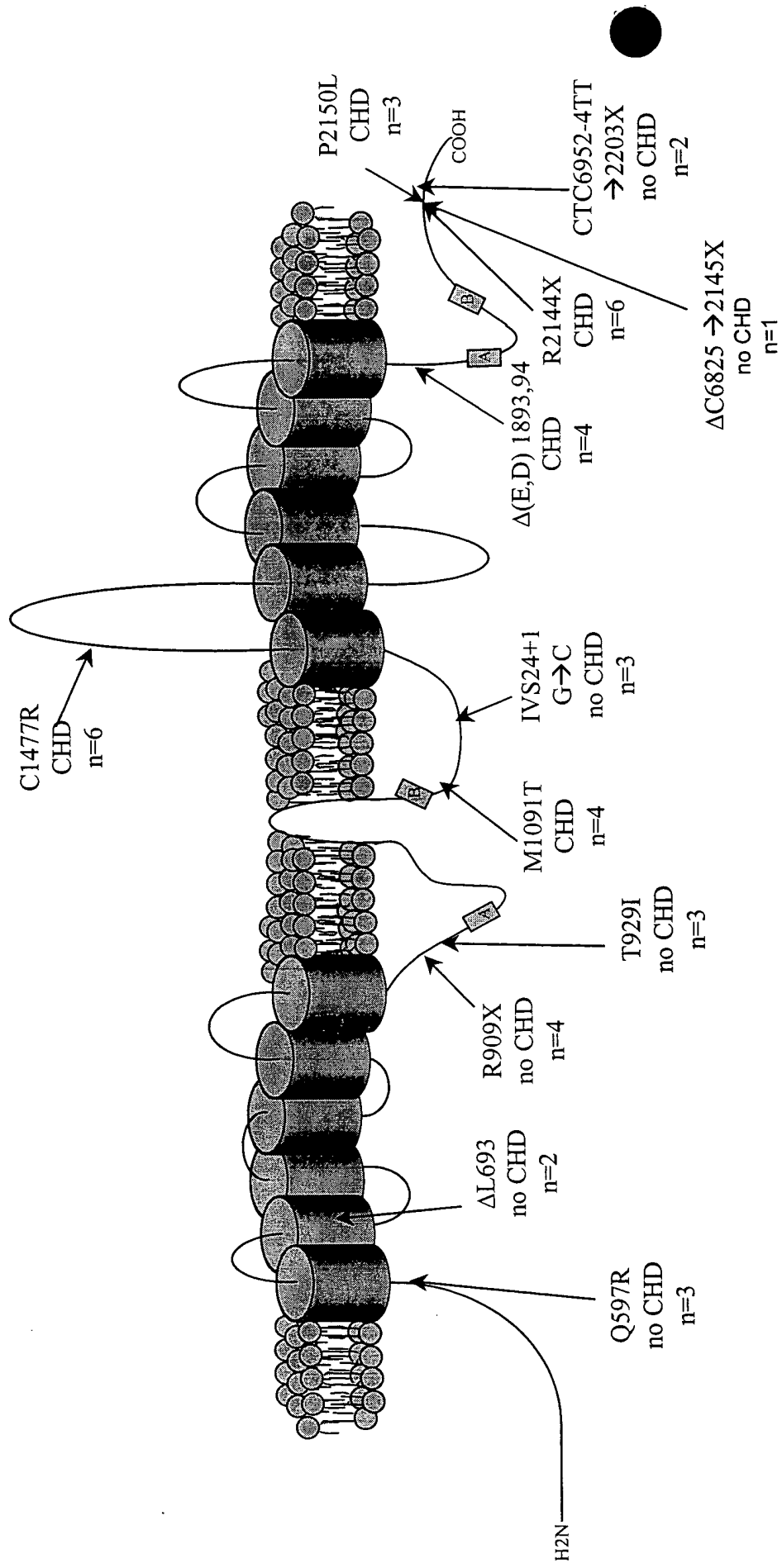
^A Calculated based on mean the age and sex specific 50th percentile levels in the LRC population

Figure 10

	Missense	Severe	P-value	Unaffected	P-Value	P-value
	Mutations	Mutations	Missense vs.	Controls	Missense vs.	Severe vs.
	(n=33)	(n=42) ^A	Severe	(n= 156)	unaffected	unaffected
TC (mmol/L)	4.78±1.30	4.30±0.95	0.08	4.71±1.07	0.76	0.02
TG (mmol/L)	1.77±2.15	1.55±1.01	0.58	1.20±1.03	0.14	0.06
HDL (mmol/L)	0.78±0.26	0.70±0.23	0.18	1.31±0.35	<0.0001	<0.0001
LDL (mmol/L)	3.19±1.10	2.90±0.91	0.23	2.84±0.87	0.10	0.73

^A for TC, TG, LDL measurements, n=41 for severe mutations, 153 for unaffected

Figure 11



any one of the following conditions is present in the family:

Figure 12A.

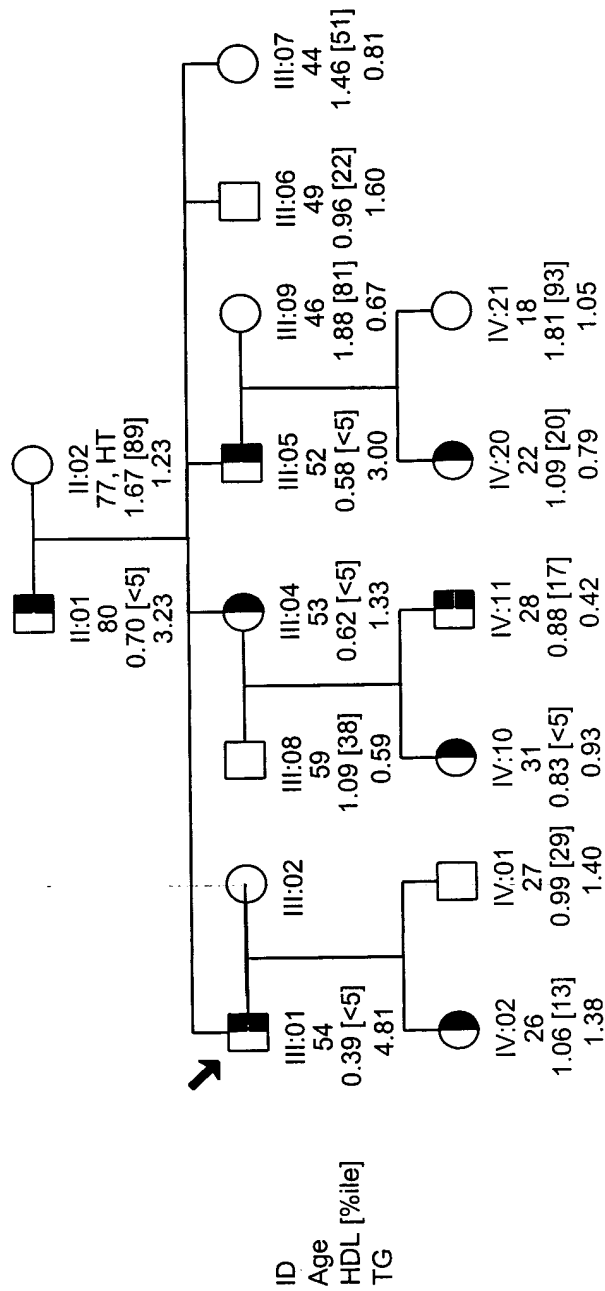


Figure 12B.

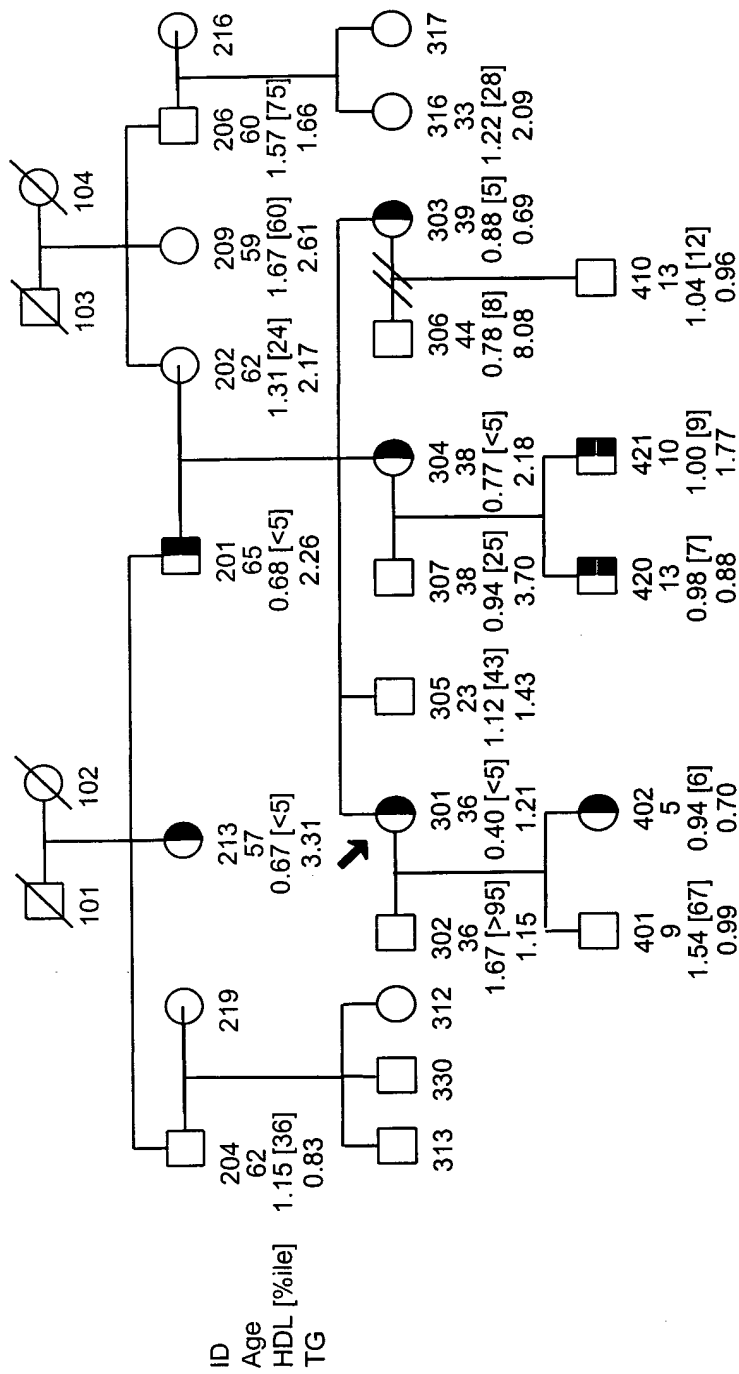


Figure 13

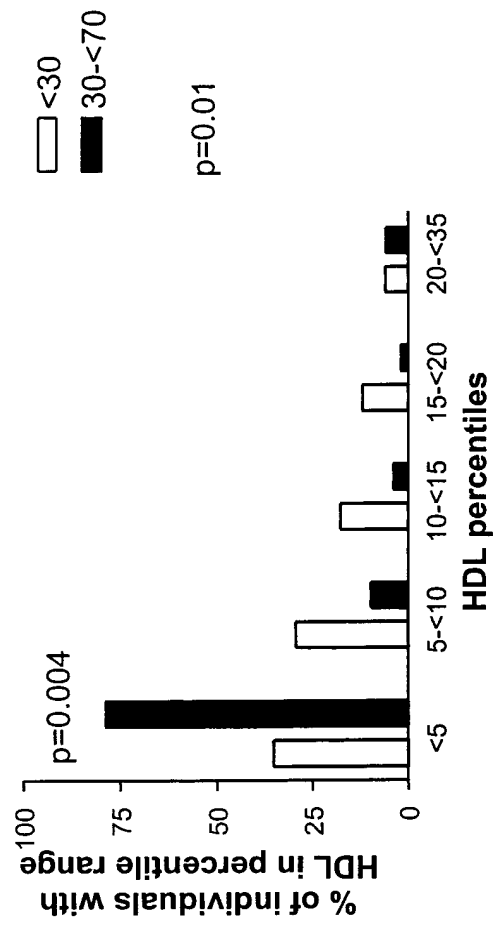


Figure 14

	Heterozygotes	Unaffected	P-value Heterozygotes
	mean±SD (n)	mean±SD (n)	vs. Unaffected
HDL (mmol/L)			
<30	0.91±0.16 (17)	1.26±0.29 (51)	<0.0001
≥30	0.66±0.24 (52)	1.32±0.36 (90)	<0.0001
Change	-0.25	+0.06	0.21
p-value <30 vs. ≥30	0.0002	0.23	
TG (mmol/L)			
<30	1.07±0.96 (16)	0.88±0.45 (51)	0.26
≥30	1.84±1.79 (52)	1.36±1.24 (87)	0.07
Change	+0.77	+0.48	0.97
p-value <30 vs. ≥30	0.03	0.001	

Figure 15A

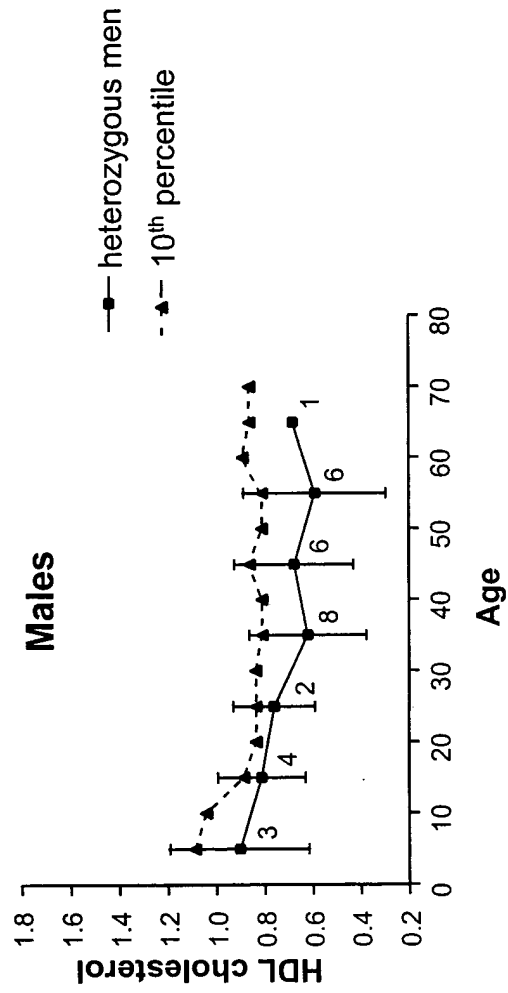
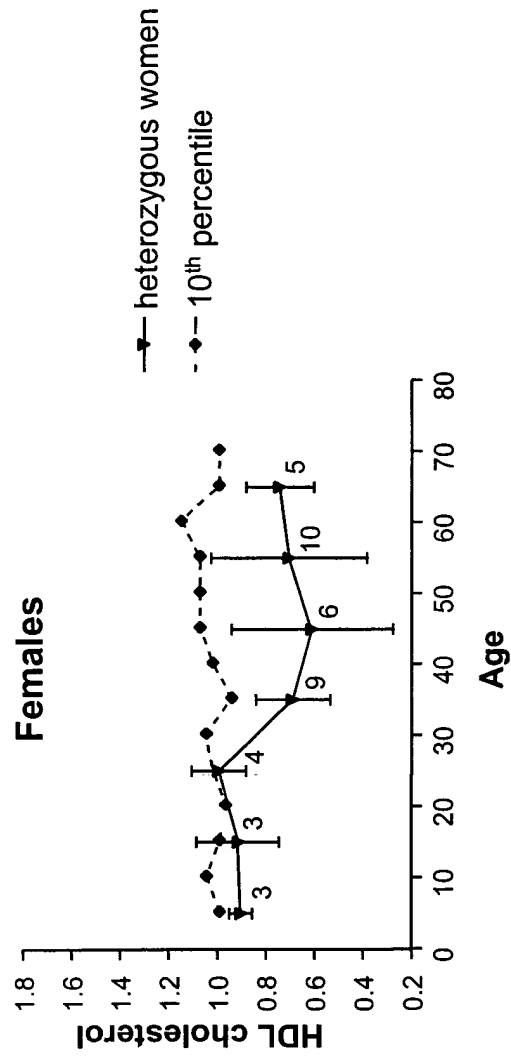


Figure 15B



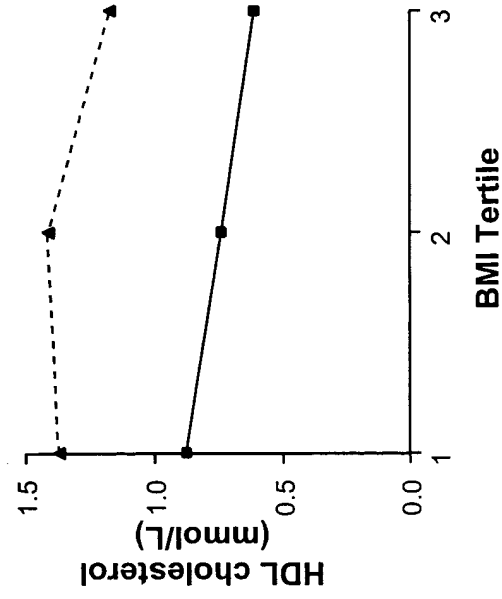


Figure 16A

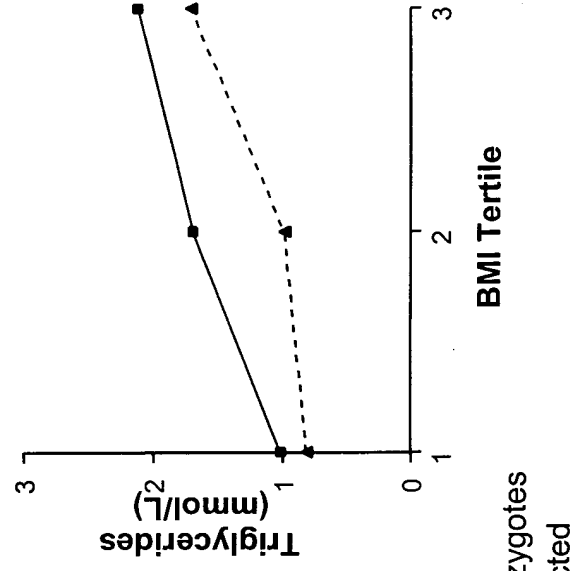


Figure 16B

Figure 17

Variant	Amount of each dNTP (uM)	Pmol each oligo	Forward oligo (5'-->3') ^a Reverse oligo (5'-->3') ^a	Annealing temp. (C)	Enzyme	Products (bp): wt "A" allele variant "B" allele	SEQ ID No:
G1051A (R219K)	187.5	20	GTATTTTGC AAGGCTACCA GTTACATTTGACAA GATTGGCTTCAGGATGTCCATGTTGGAA	60	EcoN I	177 107,70	211 212
T1591C (V399A)	200	27.5	GCTGCTGTGATGGGGTATCT ACCTCACTCACACCTGGGAA	57	Hph I	117,103,48,33 220,48,33	213 214
G2706A (V771M)	200	27.5	CAAGTGAGTGCTTGGGATTG TGCTTTTATTCAGGGACTCCA	57	BsaA I	98, 252 350	215 216
A2715C (T774P)	200	27.5	GTGATCCCAGCGTGGTGTGTCCTT GAAAGGCCAGAGGTACTCACAGCGAAGATCTTGAGGG	55	Hph I	56,69,95 56,161	217 218
G2723C (K776N)	187.5	12	TCGTTTTATTACAGGGACTCCA CAAGTGAGTGCTTGGGATTG	55	Bgl II	269,80 349	219 220
G2868A (V825I)	200	27.5	CCCATGCACTGCAGAGATTTC GCAAATTC AAAATTTCTCCAGG	57	Bsa I	149, 237 386	221 222
A3044G (I883M)	200	27.5	GAGAAGAGCCACCCCTGGTTC CAACCAGAA GAGGAT AAGGCAGGAGACATCGCTT	55	EcoR V	94,35 129	223 224
G3911C (E1172C)	200	27.5	GAGCAGTTCTGATGCTGGCCTGGGCAGCGACCA CGA TCTGCACCTCTCCTCCTCTG	55	BssS I	104, 37 141	225 226
G5155A (R1587K)	200	27.5	CAGCTTGGGAAGATTTATGACAGGACTGGACACGA ATGCCCTGCCAACTTAC	55	BssS I	114, 31 145	227 228
C5587G (S1731C)	187.5	20	GTGCAATTACGTTGTCCTGCCACACT CCATACAGCAAAAGTAGAAGGGCTAGCACA	60	Mnl I	82,35 117	229 230
G(-191)C	187.5	24	CAGCGCTTCCCGCGCGTCTTAG CCACTCACTCTCGTCCGCAATTAC	60	HgaI	287, 55, 3 342, 3	233 234

C(-17)G	187.5	18	CTGCTGAGTGACTGAACACTACATAAACAGAGGCCGGG I A CCACTCACTCTCGTCCGCAATTAC	60	Rsa I	161 124, 37	235 236
C69T	187.5	24	CAGCGCTTCCGCGCGTCTTAG CCACTCACTCTCGTCCGCAATTAC	60	BsmAI	345 310,35	237 238
C127G	187.5	24	CTGGCTTTCTGCTGAGTGAC GATCAAAAGTCCCGAAACC	60	co 0109	284, 175 459	239 240
A(-362)G	187.5	24	ACTCAGTTGTATAACCCACTGAAAAAT G AGT TTCTATAGATGTTATCATCTGGG	55	Mbo II	224, 26 134, 90, 26	241 242
A(-461)C	187.5	20	ACTCAGTTGTATAACCCACTGAAAAAT G AGT TTCTATAGATGTTATCATCTGGG	55	Hinf I	150, 100 123, 100, 27	243 244
G(-720)A	187.5	20	TCATCTAAGGCACGTTGTGG CCTCAAGCCTGGAGTGACTT	60	Hpa II	450 306, 144	245 246
G(-1027)A	187.5	20	ATGGCAAAACAGTCCTCCAAG ACCCTAGCGCTGTGTCTCTG	60	Nco I	170, 41 105, 65, 41	247 248
A(-1095)G	187.5	20	ATGGCAAAACAGTCCTCCAAG ACCCTAGCGCTGTGTCTCTG	60	MspA1 I	211 172, 39	249 250
insCCCT(-1163)	187.5	20	TGTGTGTCCTCCCTTCCATT CTTGGAGGACTGTTTGCCAT	60	Mnl I	144, 28, 11, 4 100, 48, 28, 11, 4	251 252
insG319	200	27.5	CCCCCTCGCTTTATCTTTTCAGTTAATGACCAGCCCGG ATCCCCAACTCAAAAACACACA	55	Sma I	246 210, 37	253 254
G378C	200	27.5	GCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCT G CG ATCCCCAACTCAAAAACACACA	55	Acl I	108, 41, 33, 5 141, 41, 5	255 256

^a Bold indicates mismatch in oligo to create restriction site

Figure 18

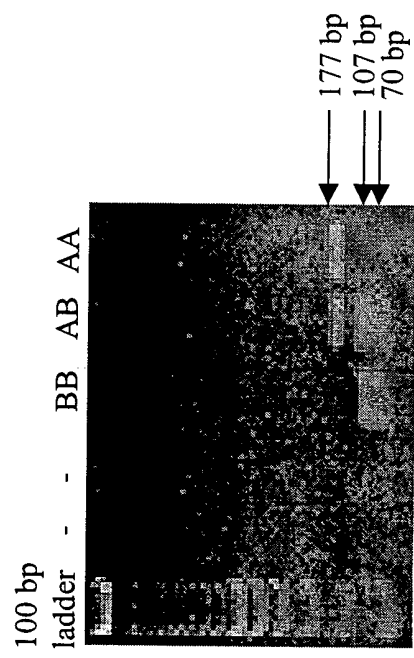


Figure 19

Nucleotide Change	Amino Acid Change	Exon	"B" allele	Frequency carrier allele (%)	N ^a	Frequency carrier allele (%)	N	Frequency carrier allele (%)	N	Frequency carrier allele (%)	N	Frequency carrier allele (%)	N	Frequency carrier allele (%)	N
Synonymous															
From Sequencing															
G869A	none	6	A	62.5	0.344	32									
C1331T	none	9	T	26.7	0.133	30									
G1343A	none	9	A	26.7	0.133	30									
T3554G	none	22	G	11.8	0.059	34									
C6842T	none	49	T	6.7	0.033	30									
Non-synonymous															
REGRESS															
G1051A	R219K	7	A	46.3	0.254	1588	39.9	0.227	546	44.1	0.247	826	47.5	0.249	686
T1591C	V399A	11	C	1.6	0.008	1098	1.2	0.006	164	1.3	0.007	150	1.0	0.005	416
G2706A	V771M	16	A	5.8	0.029	1270	6.3	0.031	318	6.3	0.031	796	3.9	0.020	406
A2715C	T774P	16	C	0.6	0.003	1250	1.3	0.006	154	1.3	0.006	156	0.6	0.003	326
G2723C	K776N	16	C	0.5	0.003	1106	0	0	178	0.6	0.003	360	1.9	0.009	318
G2868A	V825I	17	A	15.7	0.081	1364	-	-	-	-	-	-	-	-	-
A3044G	I883M	18	G	23.8	0.136	840	19.7	0.102	264	45.9	0.262	122	26.8	0.153	314
G3911C	E1172D	24	C	5.3	0.026	1288	1.3	0.007	150	6.6	0.033	822	8.2	0.041	440
G5155A	R1587K	35	A	44.3	0.259	1566	46.1	0.271	542	50.0	0.302	808	51.1	0.285	446
C5587G	S1731C	38	G	0	0	558	0	0	182	0	0	746	0	0	158
Dutch low HDL															
Dutch premature CAD															
Dutch control															
French Canadian															
G1051A	R219K	7	A	46.3	0.254	1588	39.9	0.227	546	44.1	0.247	826	47.5	0.249	686
T1591C	V399A	11	C	1.6	0.008	1098	1.2	0.006	164	1.3	0.007	150	1.0	0.005	416
G2706A	V771M	16	A	5.8	0.029	1270	6.3	0.031	318	6.3	0.031	796	3.9	0.020	406
A2715C	T774P	16	C	0.6	0.003	1250	1.3	0.006	154	1.3	0.006	156	0.6	0.003	326
G2723C	K776N	16	C	0.5	0.003	1106	0	0	178	0.6	0.003	360	1.9	0.009	318
G2868A	V825I	17	A	15.7	0.081	1364	-	-	-	-	-	-	-	-	-
A3044G	I883M	18	G	23.8	0.136	840	19.7	0.102	264	45.9	0.262	122	26.8	0.153	314
G3911C	E1172D	24	C	5.3	0.026	1288	1.3	0.007	150	6.6	0.033	822	8.2	0.041	440
G5155A	R1587K	35	A	44.3	0.259	1566	46.1	0.271	542	50.0	0.302	808	51.1	0.285	446
C5587G	S1731C	38	G	0	0	558	0	0	182	0	0	746	0	0	158
French Canadian															
G1051A	R219K	7	A	46.3	0.254	1588	39.9	0.227	546	44.1	0.247	826	47.5	0.249	686
T1591C	V399A	11	C	1.6	0.008	1098	1.2	0.006	164	1.3	0.007	150	1.0	0.005	416
G2706A	V771M	16	A	5.8	0.029	1270	6.3	0.031	318	6.3	0.031	796	3.9	0.020	406
A2715C	T774P	16	C	0.6	0.003	1250	1.3	0.006	154	1.3	0.006	156	0.6	0.003	326
G2723C	K776N	16	C	0.5	0.003	1106	0	0	178	0.6	0.003	360	1.9	0.009	318
G2868A	V825I	17	A	15.7	0.081	1364	-	-	-	-	-	-	-	-	-
A3044G	I883M	18	G	23.8	0.136	840	19.7	0.102	264	45.9	0.262	122	26.8	0.153	314
G3911C	E1172D	24	C	5.3	0.026	1288	1.3	0.007	150	6.6	0.033	822	8.2	0.041	440
G5155A	R1587K	35	A	44.3	0.259	1566	46.1	0.271	542	50.0	0.302	808	51.1	0.285	446
C5587G	S1731C	38	G	0	0	558	0	0	182	0	0	746	0	0	158
French Canadian															
G1051A	R219K	7	A	46.3	0.254	1588	39.9	0.227	546	44.1	0.247	826	47.5	0.249	686
T1591C	V399A	11	C	1.6	0.008	1098	1.2	0.006	164	1.3	0.007	150	1.0	0.005	416
G2706A	V771M	16	A	5.8	0.029	1270	6.3	0.031	318	6.3	0.031	796	3.9	0.020	406
A2715C	T774P	16	C	0.6	0.003	1250	1.3	0.006	154	1.3	0.006	156	0.6	0.003	326
G2723C	K776N	16	C	0.5	0.003	1106	0	0	178	0.6	0.003	360	1.9	0.009	318
G2868A	V825I	17	A	15.7	0.081	1364	-	-	-	-	-	-	-	-	-
A3044G	I883M	18	G	23.8	0.136	840	19.7	0.102	264	45.9	0.262	122	26.8	0.153	314
G3911C	E1172D	24	C	5.3	0.026	1288	1.3	0.007	150	6.6	0.033	822	8.2	0.041	440
G5155A	R1587K	35	A	44.3	0.259	1566	46.1	0.271	542	50.0	0.302	808	51.1	0.285	446
C5587G	S1731C	38	G	0	0	558	0	0	182	0	0	746	0	0	158
French Canadian															
G1051A	R219K	7	A	46.3	0.254	1588	39.9	0.227	546	44.1	0.247	826	47.5	0.249	686
T1591C	V399A	11	C	1.6	0.008	1098	1.2	0.006	164	1.3	0.007	150	1.0	0.005	416
G2706A	V771M	16	A	5.8	0.029	1270	6.3	0.031	318	6.3	0.031	796	3.9	0.020	406
A2715C	T774P	16	C	0.6	0.003	1250	1.3	0.006	154	1.3	0.006	156	0.6	0.003	326
G2723C	K776N	16	C	0.5	0.003	1106	0	0	178	0.6	0.003	360	1.9	0.009	318
G2868A	V825I	17	A	15.7	0.081	1364	-	-	-	-	-	-	-	-	-
A3044G	I883M	18	G	23.8	0.136	840	19.7	0.102	264	45.9	0.262	122	26.8	0.153	314
G3911C	E1172D	24	C	5.3	0.026	1288	1.3	0.007	150	6.6	0.033	822	8.2	0.041	440
G5155A	R1587K	35	A	44.3	0.259	1566	46.1	0.271	542	50.0	0.302	808	51.1	0.285	446
C5587G	S1731C	38	G	0	0	558	0	0	182	0	0	746	0	0	158
French Canadian															
G1051A	R219K	7	A	46.3	0.254	1588	39.9	0.227	546	44.1	0.247	826	47.5	0.249	686
T1591C	V399A	11	C	1.6	0.008	1098	1.2	0.006	164	1.3	0.007	150	1.0	0.005	416
G2706A	V771M	16	A	5.8	0.029	1270	6.3	0.031	318	6.3	0.031	796	3.9	0.020	406
A2715C	T774P	16	C	0.6	0.003	1250	1.3	0.006	154	1.3	0.006	156	0.6	0.003	326
G2723C	K776N	16	C	0.5	0.003	1106	0	0	178	0.6	0.003	360	1.9	0.009	318
G2868A	V825I	17	A	15.7	0.081	1364	-	-	-	-	-	-	-	-	-
A3044G	I883M	18	G	23.8	0.136	840	19.7	0.102	264	45.9	0.262	122	26.8	0.153	314
G3911C	E1172D	24	C	5.3	0.026	1288	1.3	0.007	150	6.6	0.033	822	8.2	0.041	440
G5155A	R1587K	35	A	44.3	0.259	1566	46.1	0.271	542	50.0	0.302	808	51.1	0.285	446
C5587G	S1731C	38	G	0	0	558	0	0	182	0	0	746	0	0	158
French Canadian															
G1051A	R219K	7	A	46.3	0.254	1588	39.9	0.227	546	44.1	0.247	826	47.5	0.249	686
T1591C	V399A	11	C	1.6	0.008	1098	1.2	0.006	164	1.3	0.007	150	1.0	0.005	416
G2706A	V771M	16	A	5.8	0.029	1270	6.3	0.031	318	6.3	0.031	796	3.9	0.020	406
A2715C	T774P	16	C	0.6	0.003	1250	1.3	0.006	154	1.3	0.006	156	0.6	0.003	326
G2723C	K776N	16	C	0.5	0.003	1106	0	0	178	0.6	0.003	360	1.9	0.009	318
G2868A	V825I	17	A	15.7	0.081	1364	-	-	-	-	-	-	-	-	-
A3044G	I883M	18	G	23.8	0.136	840	19.7	0.102	264	45.9	0.262	122	26.8	0.153	314
G3911C	E1172D	24	C	5.3	0.026	1288	1.3	0.007	150	6.6	0.033	822	8.2	0.041	440
G5155A	R1587K	35	A	44.3	0.259	1566	46.1	0.271	542	50.0	0.302	808	51.1	0.285	446
C5587G	S1731C	38	G	0	0	558	0	0	182	0	0	746	0	0	158
French Canadian															
G1051A	R219K	7	A	46.3	0.254	1588	39.9	0.227	546	44.1	0.247	826	47.5	0.249	686
T1591C	V399A	11	C	1.6	0.008	1098	1.2	0.006	164	1.3	0.007	150	1.0	0.005	416
G2706A	V771M	16	A	5.8	0.029	1270	6.3	0.031	318	6.3	0.031	796	3.9	0.020	406
A2715C	T774P	16	C	0.6	0.003	1250	1.3	0.006	154	1.3	0.006	156	0.6	0.003	326
G2723C	K776N	16	C	0.5	0.003	1106	0	0	178	0.6	0.003	360	1.9	0.009	318
G2868A	V825I	17	A	15.7	0.081	1364	-	-	-	-	-	-	-	-	-
A3044G	I883M	18	G	23.8	0.136	840	19.7	0.102	264	45.9	0.262	122	26.8	0.153	314
G3911C	E1172D	24	C	5.3	0.026	1288	1.3	0.007	150	6.6	0.033	822	8.2	0.041	440
G5155A	R1587K	35	A	44.3	0.259	1566	46.1	0.271	542	50.0	0.302	808	51.1	0.285	446
C5587G	S1731C	38	G	0	0	558	0	0	182	0	0	746	0	0	158
French Canadian															
G1051A	R219K	7	A	46.3	0.254	1588	39.9	0.227	546	44.1	0.247	826	47.5	0.249	686
T1591C	V399A	11	C	1.6	0.008	1098	1.2	0.006	164	1.3	0.007	150	1.0	0.005	416
G2706A	V771M	16	A	5.8	0.029	1270	6.3	0.031	318	6.3	0.031	796	3.9	0.020	406
A2715C	T774P	16	C	0.6	0.003	1250	1.3	0.006	154	1.3	0.006	156	0.6	0.003	326
G2723C	K776N	16	C	0.5	0.003	1106	0	0	178	0.6					

^a N refers to the number of alleles screened.

Figure 20

R219K	AA	AB	BB	P-value		P-value	P-value
				AA vs. AB	AA vs. BB	AA vs. AB+BB	AA vs. AB+BB
n	424	330	36				
MSD	2.70±0.37	2.77±0.37	2.78±0.40	0.01	0.22	0.005	0.005
MOD	1.73±0.35	1.81±0.35	1.85±0.35	0.002	0.05	0.001	0.001
MI before trial %(n)	48.3 (205)	47.1 (155)	33.3 (12)	0.71	0.12	0.48	0.48
events during trial %(n)	17 (71)	13 (41)	11 (4)	0.10	0.49	0.09	0.09
total events ^a (%)	65.1 (276)	59.4 (196)	44.4 (16)	0.11	0.01 ^b	0.04 ^c	0.04 ^c

^a Total events is calculated as the number of events/total number of individuals. Thus, the maximum value for this variable would be 200%, as individuals may have had events both before and during the trial.

^b Odds ratio for BB vs. AA=0.43, 95% confidence interval 0.22-0.85

^c Odds ratio for AB+BB vs. AA=0.74, 95% confidence interval 0.55-0.98

Figure 21

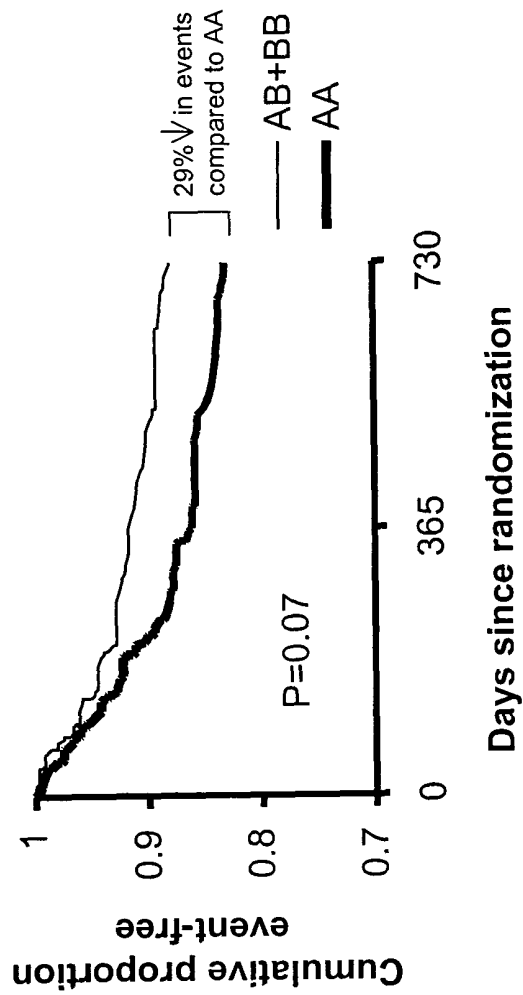


Figure 22

R219K	AA	AB	BB	P-value		
				AA vs. AB	AA vs. BB	AA vs. AB+BB
n	424	330	36			
Age	57±8	55±8	57±7	0.0007	1	0.03
BMI	25.8±2.6	26.3±2.7	25.5±2.3	0.01	0.50	0.09
Total Cholesterol	6.02±0.86	6.07±0.89	5.89±0.85	0.44	0.38	0.60
HDL Cholesterol	0.92±0.22	0.93±0.23	0.92±0.20	0.54	1	0.81
LDL Cholesterol	4.27±0.75	4.35±0.83	4.33±0.82	0.17	0.65	0.19
Triglycerides	1.84±0.77	1.78±0.78	1.42±0.49	0.29	0.001	0.08

Downloaded from <https://www.cambridge.org/core>.
 University of Cambridge, on 01 Jun 2018 at 10:00:00, subject to the Cambridge Core terms of use, available at <https://www.cambridge.org/core/terms>.
<https://doi.org/10.1017/S0007122618000000>

Figure 23

		< median		> median		P-value		P-value		P-value	
		n	mean±SD	n	mean±SD	< vs. >median		AB+BB vs. AA		AB+BB vs. AA	
								< median		> median	
AB+BB											
Total Cholesterol	193	6.22±0.91	172	5.87±0.82	0.0001	0.22	0.43				
HDL cholesterol	192	0.91±0.22	171	0.94±0.23	0.21	0.12	0.37				
LDL Cholesterol	192	4.49±0.84	171	4.19±0.78	0.0005	0.03	0.57				
Triglycerides	193	1.82±0.79	172	1.65±0.72	0.03	0.02	0.85				
MSD	193	2.79±0.37	171	2.75±0.37	0.30	0.18	0.01				
MOD	193	1.83±0.36	171	1.78±0.34	0.18	0.09	0.006				
AA											
Total Cholesterol	207	6.11±0.86	217	5.94±0.84	0.04						
HDL cholesterol	206	0.88±0.20	214	0.96±0.24	0.0002						
LDL Cholesterol	205	4.32±0.77	214	4.23±0.72	0.22						
Triglycerides	206	2.02±0.82	217	1.67±0.67	<0.0001						
MSD	205	2.75±0.36	217	2.65±0.38	0.006						
MOD	205	1.77±0.34	217	1.69±0.35	0.04						

Figure 24

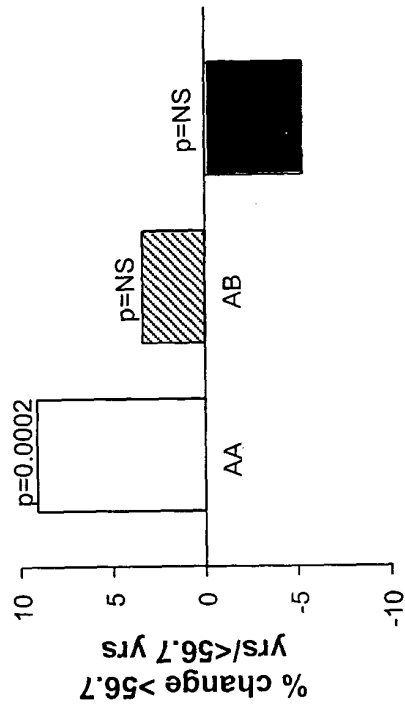


Figure 25A

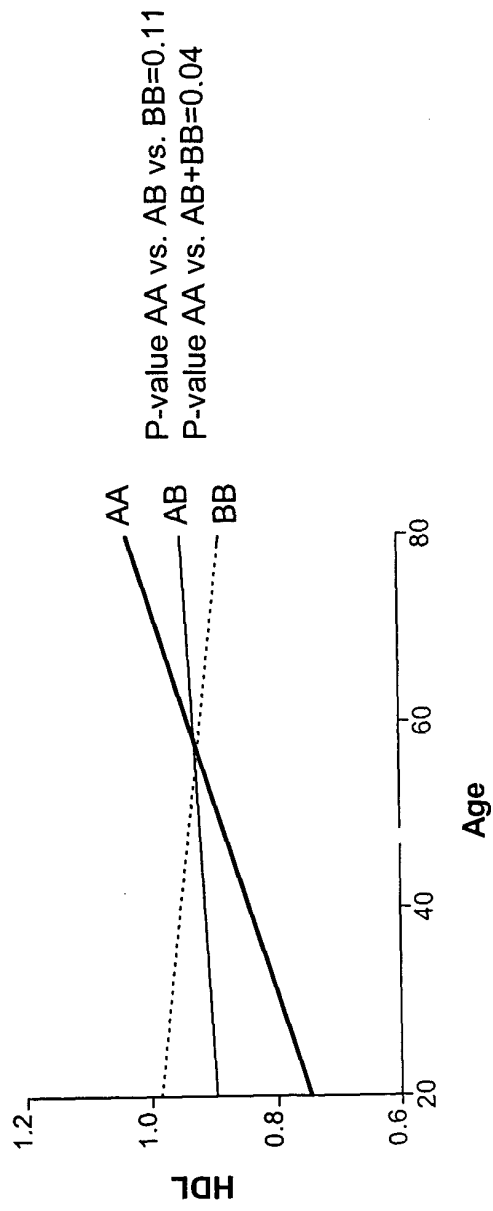


Figure 25B

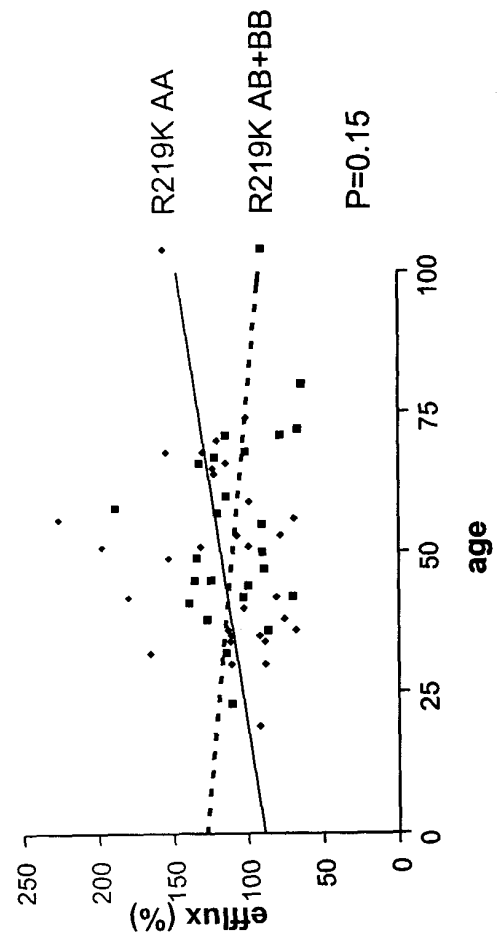


Figure 26A

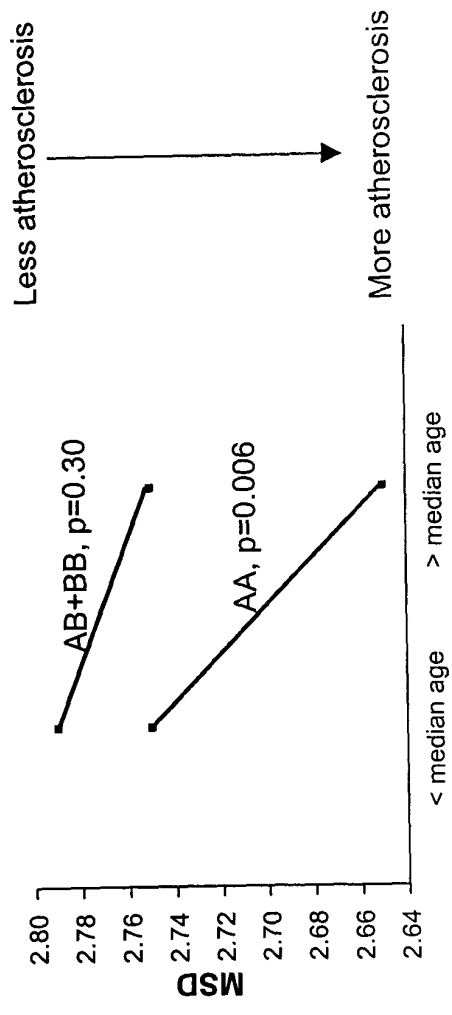


Figure 26A

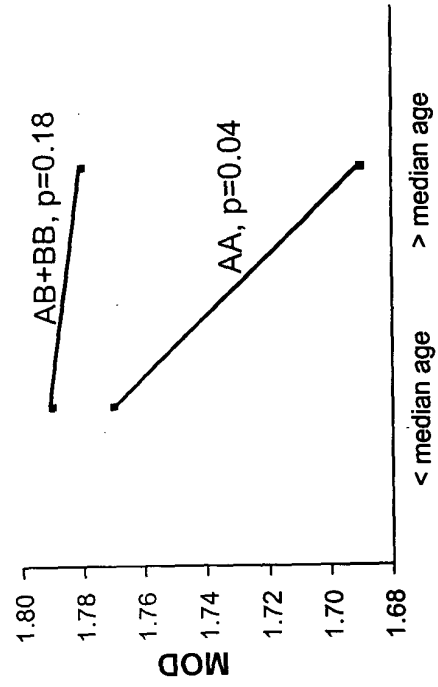


Figure 27

	South African Black ^a	Cantonese	Dutch ^b	P-value South African vs. Dutch	P-value Cantonese vs. Dutch
AA % (n)	1.3 (1)	32.7 (33)	52.5 (180)		
AB % (n)	50.7 (38)	55.4 (56)	45.2 (155)	<0.0001	<0.0001
BB % (n)	48.0 (36)	11.9 (12)	2.3 (8)		
n	75	101	343		
carrier freq.	98.67	67.33	47.52	<0.0001	0.0005
allele freq.	0.733	0.396	0.249	<0.0001	<0.0001

^a Not consistent with Hardy Weinberg equilibrium (p=0.01)

^b Not consistent with Hardy-Weinberg equilibrium (p<0.001)